



Masterarbeit

zum Thema

Akzeptanz von Fernbushaltestellen in Routennähe mit direkter Anbindung an das Netz des ÖPNV – Eine Conjoint - Analyse

vorgelegt von:	Jana-Christina Hambrock
Matrikel-Nr.:	3750149
Studiengang:	Verkehrswirtschaft
geboren am:	30.08.1992 in Hannover
Verantwortlicher	
Hochschullehrer:	Prof. Dr. rer. pol. Ostap Okhrin
Betreuer:	Dr. Kathrin Kormoll

Dresden, den *08.08.2017*

Abstract

Seit der Liberalisierung im Jahr 2013 wächst die Anzahl der Fernbusfahrgäste immer weiter an. Durch ein vielfältiges Streckenangebot und günstige Ticketpreise erfreut sich der Fernbus immer größerer Beliebtheit und ist zu einer guten und umweltfreundlichen Alternative für Fernreisen geworden. Allerdings kann die Reisezeit bei den Angeboten der Konkurrenten wenig mithalten, was vor allem an der Lage der Haltestellen liegt. In dieser Arbeit wird mittels auswahlbasierter Conjoint-Analyse die Kundensicht zu Haltestellen in der Stadt und Haltestellen in Routennähe untersucht. Dabei nehmen vor allem Reisezeit und Reisekosten Einfluss auf die Entscheidung. Sozioökonomische Merkmale von Kunden spielen nur eine untergeordnete Rolle. Es kann gezeigt werden, dass Haltestellen in Routennähe für Fernbusse eine ernstzunehmende Alternative darstellen, ungeachtet von der Länge der Strecke, solange eine gute Anbindung an das Netz des öffentlichen Nahverkehrs gegeben ist.

Since the liberalisation in 2013, the amount of intercity bus users is continuously increasing. The concept of offering a large variety in routes as well as attractive pricing, helps intercity buses to a great popularity. Moreover, this option of traveling impresses with its eco-friendliness. Due to the location of the bus stops, the traveling time in intercity busses cannot keep up with the one of comparable competitors. Within this thesis, a selective conjoint analysis is conducted. The goal is to gain an insight on the option of potential customers towards bus stops within the city compared to their view of bus stops close to the traveling route. The decision is primarily based on travel costs and travel time whereas socioeconomic factors are only considered briefly. Concluding, bus stops close to the traveling route are an alternative choice for as long as they are located within reach of local public transportation.

Inhaltsverzeichnis

ABBILDUNGSVERZEICHNIS.....	III
TABELLENVERZEICHNIS	IV
SYMBOLVERZEICHNIS.....	VI
ABKÜRZUNGSVERZEICHNIS.....	VIII
1. EINLEITUNG	9
1.1 MOTIVATION	9
1.2 ORGANISATION DIESER ARBEIT	9
2. DER FERNBUSVERKEHR	10
2.1 ALLGEMEIN	10
2.2 PERSPEKTIVEN UND ENTWICKLUNGSMÖGLICHKEITEN	11
2.3 AKTUELLE STUDIEN.....	12
3. METHODISCHES VORGEHEN	17
3.1 DATENSPEZIFIKATION.....	17
3.1.1 Fragebogen.....	17
3.1.2 Daten	19
3.2 METHODEN	19
3.2.1 Allgemein.....	19
3.2.2 Auswahlbasierte Conjoint-Analyse.....	21
3.2.3 Logistische Regression	28
4. ANALYSEN	32
4.1 ZIELGRUPPENANALYSE.....	32
4.1.1 Zielgruppen.....	33
4.1.2 Merkmalsanalyse	33
4.2 KORRELATIONSANALYSE.....	44
4.3 AUSWAHLBASIERTE CONJOINT-ANALYSE	45
4.3.1 Lange Strecken	46
4.3.2 Kurze Strecken.....	54
4.3.3 Lange und kurzen Strecken zusammen	59
4.3.4 Ergebnis der auswahlbasierten Conjoint-Analyse.....	65
4.4 LOGISTISCHE REGRESSION.....	66
4.4.1 Lange Strecken	67



4.4.2	<i>Kurze Strecken</i>	71
4.4.3	<i>Ergebnis der logistischen Regression</i>	74
4.5	ERGEBNIS DER ANALYSEN	75
5.	KRITISCHE WÜRDIGUNG	76
6.	ZUSAMMENFASSUNG UND AUSBLICK	77
	QUELLENVERZEICHNIS	79
	ANHANG	82

Abbildungsverzeichnis

Abbildung 1: Methoden zur Präferenzmessung (Fabian (2005), S. 116).....	20
Abbildung 2: Relative Häufigkeit des Merkmals Geschlecht (e.D.).....	34
Abbildung 3: Relative Häufigkeit des Merkmals Beruf (e.D.)	38
Abbildung 4: Relative Häufigkeit des Merkmals Höchster Schulabschluss (e.D.).....	41
Abbildung 5: Relative Häufigkeit des Merkmals Höchste Berufsausbildung (e.D.)	42
Abbildung 6: Relative Häufigkeit der PKW Verfügbarkeit (e.D.).....	43
Abbildung 7: Geschätzte Teilnutzenwerte (e.D.).....	48
Abbildung 8: Relative Wichtigkeit der Eigenschaften der langen Strecken (e.D.).....	49
Abbildung 9: Gesamtnutzenwerte für die langen Strecken (e.D.).....	50
Abbildung 10: Gesamtnutzenwerte der langen Strecke mit Aufspaltung der Reisezeit (e.D.)	53
Abbildung 11: Gesamtnutzenwerte der Varianten für kurze Strecken (e.D.)	56
Abbildung 12: Gesamtnutzenwerte der kurzen Strecke mit Aufspaltung der Reisezeit (e.D.)	58
Abbildung 13: Relative Wichtigkeit der Variablen bei beiden Strecken (e.D.).....	60
Abbildung 14: Gesamtnutzenwerte der Varianten für beide Strecken (e.D.).....	61
Abbildung 15: Gesamtnutzenwerte der Varianten für beide Strecken mit expliziten Werten (e.D.)....	64
Abbildung 16: Wahrscheinlichkeitsverteilung lange Strecken (e.D.)	69
Abbildung 17: Ausreißer lange Strecken (e.D.).....	70
Abbildung 18: Wahrscheinlichkeitsverteilung kurze Strecken (e.D.).....	72
Abbildung 19: Ausreißer kurze Strecken (e.D.).....	74
Abbildung 20: Wahrscheinlichkeitsverteilung für beide Strecken (e.D.)	75

Tabellenverzeichnis

Tabelle 1: Vor- und Nachteile der Lage von Haltestellen (e.D. nach Selzer (2016), S. 17).....	13
Tabelle 2: Anteile der verwendeten Verkehrsmittel zur An- und Abreise zur Haltestelle (e.D. nach FlixBus (2014))	15
Tabelle 3: Wichtigkeit der Standortmerkmale von Fernbushaltestellen (e.D. nach FlixBus (2014))....	16
Tabelle 4: Relative Häufigkeiten des Merkmals Alter	35
Tabelle 5: Lagemaße für das Merkmal Alter.....	37
Tabelle 6: Relative Häufigkeiten des Merkmals Einkommen (e.D.).....	39
Tabelle 7: Lagemaße für das Merkmal Einkommen	40
Tabelle 8: Relative Häufigkeiten des Merkmals Fernbusnutzung (e.D.)	44
Tabelle 9: Chi ² -Unabhängigkeitstest von Alter und Zielgruppenzugehörigkeit.....	45
Tabelle 10: Darstellung der Varianten bei langen Strecken	46
Tabelle 11: Ermittlung Startwerte der langen Strecken.....	47
Tabelle 12: Schätzwerte der Teilnutzen nach ML-Schätzung der langen Strecken	47
Tabelle 13: Güteprüfung des Modells (Lange Strecken).....	50
Tabelle 14: Prüfung der Koeffizienten der langen Strecken.....	51
Tabelle 15: Teilnutzen lange Strecke mit Aufspaltung der Zeit.....	52
Tabelle 16: Darstellung der Varianten bei kurzen Strecken	54
Tabelle 17: Ermittlung Startwerte und Schätzwerte nach ML-Schätzung der kurzen Strecken.....	55
Tabelle 18: Relative Wichtigkeit der Variablen	55
Tabelle 19: Güteprüfung des Modells (kurze Strecken).....	56
Tabelle 20: Prüfung der Koeffizienten der kurzen Strecken	57
Tabelle 21: Teilnutzenwerte der kurzen Strecke mit Aufspaltung der Reisezeit	58

Tabelle 22: Ermittlung Startwerte und Schätzwerte nach ML-Schätzung beider Strecken	60
Tabelle 23: Güteprüfung des Modells (lange und kurze Strecken).....	62
Tabelle 24: Prüfung der Koeffizienten für beide Strecken.....	62
Tabelle 25: Rel. Wichtigkeit der Variablen beider Strecken mit expliziten Werten.....	63
Tabelle 26: Ergebnisse logistische Regression der langen Strecken.....	67
Tabelle 27: Odds Ratio der langen Strecken.....	68
Tabelle 28: Klassifikationstabelle lange Strecken.....	69
Tabelle 29: Ergebnisse logistische Regression der kurzen Strecken.....	71
Tabelle 30: Odds Ratio der kurzen Strecken.....	72
Tabelle 31: Klassifikationstabelle kurze Strecken	73

Symbolverzeichnis

a	Parameter der logistischen Regression
b	Teilnutzen
β	geschätzter Parameter der logistischen Regression
C	Klasse
d_c	Klassenbreite
eig	Variablen der logistischen Regression
f_c	relative Häufigkeit der Klasse c
F_C	kumulierte Häufigkeit der Klasse c
g	Auswahlentscheidungen
h	absolute Häufigkeit
i	Indize
j	Eigenschaft
k	Alternative
L	Likelihood-Wert
L_0	Likelihood-Wert des Nullmodells
L_V	Likelihood-Wert des vollständigen Modells
LL	LogLikelihood-Wert
LL_0	LogLikelihood-Wert des Nullmodells
LLR	LogLikelihood-Ratio
m	Ausprägung
n	Anzahl der Personen
n_a	Anzahl der Auswahlentscheidungen
n_{FBN}	Anzahl der Fernbusnutzer
n_{NN}	Anzahl der Nichtnutzer
o	Fehlerterm

p	Indize
$prob$	Wahrscheinlichkeit
ρ	Rationalitätsparameter
Π	Produkt
r	Situation
R_{XYI}	Rangkorrelationskoeffizient zwischen der Variable X und Variable YI
R^2	Bestimmtheitsmaß
R^2_M	McFadden- R^2
$RESID$	Residuen
st	Stimuli
s^2	Varianz
s	Standardabweichung
Σ	Summe
T	Auswahlsituation
u	Nutzen
V	Variationskoeffizient
w	Dummyvariable
x_u	Klassenuntergrenze
x_o	Klassenobergrenze
x_C^*	Klassenmitte
\bar{x}	Mittelwert
x_C^u	Untergrenze der Klasse
x_C^o	Obergrenze der Klasse
x_{Med}	Median
χ^2	Chi ² -Verteilung
y	abhängige Variable
z	Logit

ZResid standardisierte Residuen

Abkürzungsverzeichnis

<i>ADAC</i>	Allgemeiner Deutscher Automobil Club
<i>AIC</i>	Akaike Informations Criterion
<i>CBCA</i>	Choice based Conjoint-Analyse
<i>CS</i>	Choice Set
<i>DB</i>	Deutsche Bahn
<i>e.D.</i>	eigene Darstellung
<i>FBN</i>	Fernbusnutzer
<i>HB</i>	Hierarchical Bayes
<i>IC</i>	InterCity
<i>LC</i>	Latent Class
<i>ML</i>	Maximum-Likelihood
<i>NE</i>	Nutzeneinheiten
<i>NN</i>	Nicht-Nutzer
<i>ÖPNV</i>	öffentlicher Personennahverkehr
<i>ÖV</i>	öffentlicher Verkehr
<i>PBefG</i>	PersonenBeförderungGesetz
<i>PKW</i>	Personenkraftwagen
<i>ZOB</i>	zentraler Omnibus Bahnhof

1. Einleitung

1.1 Motivation

Seit der Fernbusliberalisierung im Jahr 2013 wächst der Markt der Fernreisebusse stetig an. Ob auf kurzen oder langen Strecken, mittlerweile findet sich für jede Reise eine Route, die von Fernbussen befahren wird. Durch die günstigen Ticketpreise ist der Bus eine gute Alternative zum Schienenverkehr und so vor allem bei jungen Menschen beliebt. Allerdings muss bei einer Fahrt im Fernbus, besonders auf längeren Strecken mit mehreren Zwischenhalten, sehr viel mehr Zeit eingeplant werden, als bei anderen Verkehrsmitteln. Dies liegt zum großen Teil daran, dass sich die Haltestellen in den Stadtzentren befinden, und so die Busse neben der normalen Autobahnstrecke auch noch den Stadtverkehr überwinden müssen. In der folgenden Arbeit soll eine Conjoint-Analyse die Frage zur Akzeptanz von Fernbushaltestellen in Routennähe beantworten. Zusätzlich wird der „typische“ Fernbusnutzer charakterisiert und der Einfluss von personenbezogenen Merkmalen auf die Auswahlentscheidung untersucht. Es zeigt sich, dass Kunden vor allem auf langen Strecken eine Haltestelle in Routennähe bevorzugen, da so erheblich viel Zeit eingespart werden kann. Ein wichtiger Aspekt dabei ist eine gute ÖPNV-Anbindung der Haltestelle, um von den etwas außerhalb gelegenen Plätzen zum Zielort zu gelangen. Auf kurzen Strecken, die keine Zwischenhalte anfahren, werden die Haltestellen in der Stadt bevorzugt, da hier die komplexe Reisezeit geringer ist. Des Weiteren konnte festgestellt werden, dass Fernbusnutzer zwischen 18 und 30 Jahre alt sind und ein geringes Einkommen haben, da sie meist noch studieren und somit sehr gebildet sind. Ein eigener PKW ist nicht vorhanden. In einer logistischen Regression zeigt sich, dass Geschlecht, Berufsstatus, Einkommen und PKW-Verfügbarkeit einen geringen Einfluss auf die Wahlentscheidung der bevorzugten Haltestellenlage besitzen.

1.2 Organisation dieser Arbeit

In Kapitel 2 werden zunächst allgemeine Informationen zum Thema Fernbusverkehr vorgestellt. Dabei wird es unter anderem einen kleinen Überblick über die Entwicklung nach der Liberalisierung geben, die Perspektiven und Möglichkeiten des Fernbusmarktes für die nächsten Jahren werden erläutert und anhand von verschiedenen Studien sollen die Kundenanforde-

rungen an die Erreichbarkeit, Lage und Ausstattung von Fernbushaltestellen vorgestellt werden. In Kapitel 3 wird das methodische Vorgehen der vorliegenden Arbeit präsentiert. Dabei wird der verwendete Fragebogen der Untersuchung erläutert und die erhobenen Daten werden kurz beschrieben. Im Anschluss daran werden die theoretischen Grundlagen für die Conjoint-Analyse und die logistische Regression gelegt. In Kapitel 4 werden zunächst die sozioökonomischen Merkmale der Probanden analysiert, um den „typischen“ Fernbusnutzer zu identifizieren. Die anschließende Conjoint-Analyse soll die Frage nach der Akzeptanz der Haltestellenlage von Kunden beantworten. Dabei werden die Analysen für lange und kurze Strecken voneinander getrennt und anschließend auch zusammen betrachtet. Abschließend werden die Auswahlentscheidungen der Probanden anhand ihrer personenspezifischen Merkmale durch eine logistische Regression untersucht. Im fünften Kapitel wird die Analyse kritisch hinterfragt und eventuelle Probleme oder Verbesserungen werden aufgezeigt.

2. Der Fernbusverkehr

Im Folgenden Kapitel soll es einen ersten Überblick über den Fernbusverkehr geben. Dazu werden erst allgemeine Grundlagen in der Geschichte und über die Thematik gelegt. Anschließend sollen die Chancen und Perspektiven des Fernbusmarktes aufgezeigt werden. Zum Schluss werden aktuelle Studien vorgestellt, die sich mit den Anforderungen an Standort und Erreichbarkeit von Haltestellen für Fernbusse beschäftigen.

2.1 Allgemein

Vor der Liberalisierung zum 01.01.2013 mussten neue Buslinien im Fernverkehr genehmigt werden, um den Schienenverkehr nicht zu gefährden. Diese Marktzutrittsschranke für Fernbusse bestand seit dem Jahr 1935. Dabei gab es keine Genehmigung, wenn für die vorgesehene Strecke bereits eine Eisenbahnverbindung bestand (PBefG §13). Ausnahme hierbei waren internationale Verbindungen, welche allerdings ein innerdeutsches Beförderungsverbot hatten. (Böckers (2015), S. 76/77) Die Voraussetzungen für eine Zulassung waren sehr hart und die Preise wurden kontrolliert, sodass keine Gefahr für die Bahnen bestand (Mäertens (2012), S. 8). Eine Ausnahme war damals das Unternehmen BerlinLinienBus, welches ein Tochterunternehmen der Deutschen Bahn (DB) ist. Seit Inkrafttreten der neuen Novelle des Personen-Beförderungsgesetzes (PBefG) zum 1. Januar 2013 können Fernbusse ungehindert in

Deutschland fahren, solange die Distanz mehr als 50 Kilometer beträgt oder die Fahrt länger als eine Stunde dauert (Moonroc (2013), S. 2). Seit diesem Tag sind viele Unternehmen in das Fernbusgeschäft eingestiegen. Laut Statistischem Bundesamt fuhren im Jahr 2013 ca. 8,2 Millionen Personen mit Fernbussen und langfristig kann von bis zu 25 Millionen Nutzern im Jahr ausgegangen werden (Böckers (2015), S. 77). Laut Ratzenberger (2015) haben sich die Buslinien bereits bis zum September 2014 verdreifacht und zwischen 30 und 44 Prozent der Nutzer sind vorher Bahnkunden gewesen. Allerdings konnte zwischen September 2015 und März 2016 ein Rückgang von Linien verzeichnet werden. Es kann somit eine starke Dynamik im Fernbusmarkt festgestellt werden. Ende 2014 waren ca. 20 verschiedene Fernbusunternehmen mit mehr als 7.000 Fahrten in der Woche auf dem Markt aktiv. Zu den größten Anbietern zählen Unternehmen wie MeinFernbus, FlixBus, ADAC Postbus sowie die Tochterunternehmen der DB BerlinLinienBus und IC Bus, wobei MeinFernbus und FlixBus seit dem Jahr 2015 ein gemeinsames Unternehmen sind. Zusammen kontrollieren sie fast 75 Prozent des deutschen Fernbusmarktes. (Böckers (2015), S. 77/78) Zwischenzeitlich drängten sich auch viele kleinere Unternehmen auf den Markt, sind aber nach kurzer Zeit durch den enormen Preiskampf zwischen den Anbietern wieder verdrängt worden (Knorr (2016), S. 327-329). Mittlerweile hat sich der Fernbus zu einer starken Konkurrenz für den Schienenverkehr entwickelt. Allerdings ist der Markt nach vier Jahren noch relativ neu und die zukünftigen Auswirkungen können noch nicht genau abgeschätzt werden (Böckers (2015), S. 87-89). Die Perspektiven und Entwicklungsmöglichkeiten werden im nachfolgenden Kapitel 2.2 näher betrachtet.

2.2 Perspektiven und Entwicklungsmöglichkeiten

Laut dem Moonroc Institute of Economic Research lagen die Schätzungen für eine potentielle Marktgröße des Fernbusses im Jahr 2013 zwischen 0,3 und 5 Milliarden Euro. Dabei wird ein Wachstum des Fernbusmarktes vom zunehmenden Mobilitätsbedarf, den steigenden Mobilitätskosten und eines höheren Umweltbewusstseins der Menschen begünstigt (Moonroc (2013), S. 2). Ein Fernbus verbraucht pro Passagier weniger Energie als ein Auto und ist daher ein sehr effizientes Verkehrsmittel, welches zur Zeit von steigendem Umweltbewusstsein gute Chancen besitzt. Erfahrungen aus den USA oder Großbritannien zeigen, dass die anfänglich große Anzahl an Unternehmen sinken wird, der Fernbusmarkt jedoch weiterhin eine wichtige Alternative in der Mobilitätsfrage bleiben wird (Augustin (2014), S. 245). Die Chancen für

die Entwicklung stehen gut, da der Fernbus im Gegensatz zu anderen Verkehrsmitteln günstige Preise und eine umweltfreundliche Alternative bietet. Die Erfahrungen aus Nachbarmärkten wie Schweden oder Großbritannien lassen dies ebenfalls vermuten. (Moonroc (2013), S. 4/5) Dennoch kann nicht davon ausgegangen werden, dass die Preise für Tickets so niedrig bleiben. Da allerdings der Preis wettbewerbsbestimmend ist, sind ein deutschlandweites Streckennetz und eine erhöhte Auslastung notwendig. Laut dem Bundesministerium für Verkehr und digitale Infrastruktur kann mit einer steigenden Anzahl an Fahrgästen gerechnet werden. Der Fernbusmarkt besitzt ein großes Potential, welches nur durch fehlendes qualifiziertes Personal, der Umsetzung der Barrierefreiheit oder durch schlechte Umsetzung von Haltestellen gehemmt werden könnte. Vorstellbar für die Zukunft sind Nachtbuslinien, Expresslinien oder Linien, die zu touristischen Gebieten oder ins Ausland fahren. (Selzer (2016), S. 11-13) Es ist wahrscheinlich, dass sich in Deutschland ähnlich wie in Schweden ein Oligopol aus drei bis vier Anbietern entwickeln wird, welche den Markt kontrollieren werden (Selzer (2016), S. 20). Eine weitere denkbare Maßnahme für die Zukunft ist die Fernbusmaut. Zur jetzigen Zeit besitzen die Fernbusse einen Wettbewerbsvorteil gegenüber dem Schienenpersonenverkehr, da jede befahrene Trasse Geld kostet. Für Fernbusse gilt so eine Regelung noch nicht. Eine Fernbusmaut ist ökonomisch sinnvoll, da diese den intermodalen Wettbewerb fördert und eventuelle zusätzliche Abnutzung der Autobahnen refinanziert. (Böckers (2015), S. 88).

2.3 Aktuelle Studien

Im folgenden Teil sollen einige aktuelle Studien zu Anforderungen an die Erreichbarkeit, Lage und Ausstattung von Fernbushaltestellen vorgestellt werden.

In einer Studie aus dem Jahr 2016 von Selzer, Kruse, Wilde und Lanzendorf werden die Anforderungen an die Erreichbarkeit und Lage von Fernbusterminals mit Hilfe einer Umfrage in Frankfurt am Main untersucht. Des Weiteren sollen verschiedene städtebauliche und verkehrliche Handlungsoptionen vorgestellt werden. Seit Beginn der Fernbusliberalisierung konnte eine verbesserte Erreichbarkeit von mittelgroßen Städten festgestellt werden. Zusätzlich wird die Haltestelleninfrastruktur näher betrachtet. Dabei können sehr unterschiedliche Ausbauweisen von Fernbushaltestellen festgestellt werden. Zum einen gibt es in einigen Städten voll ausgebaute Zentrale Omnibus Bahnhöfe (ZOB) mit vollständigem Terminal. Zum anderen sind in einigen Städten ZOBs vorhanden, die schon vor der Liberalisierung von Fern-

verkehr genutzt wurden. Andere Städte nutzen einen regionalen Busbahnhof oder auch nur Straßenrandhaltestellen. Diese Einrichtungen besaßen vor dem 1. Januar 2013 nur eine geringe Auslastung. (Selzer (2016), S. 13-16) Die Forschungsgesellschaft für Straßen- und Verkehrswesen (FGSV) unterscheidet zwischen vier verschiedenen Typen von Haltestellen, die sowohl von Funktion als auch von Fahrgast- und Fahrzeugaufkommen abhängen. Diese reichen von voll ausgebauten Busterminals mit Werkstätten und Warteräumen bis hin zu Straßenrandhaltestellen. Ein weiterer wichtiger Punkt für Haltestellen ist die Lage innerhalb der Städte. Hier wird zwischen zentral und dezentral unterschieden. Dabei sind zentrale Haltestellen die, die in der Innenstadt mit sehr guter ÖPNV-Anbindung und meist in der Nähe von Bahnhöfen liegen. Diese werden von Kunden bevorzugt. Dezentrale liegen meist in Autobahnnähe und haben betriebswirtschaftliche Vorteile für die Unternehmen, da sie die Reisezeitverluste vermindern, die bei der Fahrt in die Innenstadt anfallen. (Selzer (2016), S. 16/17) Eine Zusammenfassung der Vor- und Nachteile ist in Tabelle 1 abgebildet.

Lage	Vorteile	Nachteile
Zentrale Lage	+ Sehr guter ÖPNV-Anschluss	- Abgase und Lärm
	+ Innenstadtinfrastruktur ist vorhanden	- Flächenbedarf
	+ nahegelegene Versorgungsangebote	- Zeitverluste durch den Innenstadtverkehr
Dezentrale Lage	+ Geringe Zeitverluste durch Nähe zur Autobahn	- Längere ÖPNV-Anbindung
	+ Geringe Verkehrsbelastung	- Unter Umständen nicht so dichtes ÖPNV-Angebot
	+ Flächen verfügbar und günstiger	

Tabelle 1: Vor- und Nachteile der Lage von Haltestellen (e.D. nach Selzer (2016), S. 17)

Von 112 Haltestellen befinden sich zurzeit 91 in der Innenstadt an nahegelegenen Bahnhöfen. (Selzer (2016), S. 17-19) Die Studie untersucht mittels Kundenbefragung in Frankfurt am Main die Anforderungen an Erreichbarkeit und Lage. Dabei wird zum einen der „typische“ Fernbusnutzer beschrieben und zum anderen nach der idealen Lage aus Kundensicht gefragt. Laut Umfrage sind 54 Prozent der Nutzer Frauen und 46 Prozent Männer. Der Groß-

teil der Nutzer ist jünger als 35 Jahre und studiert oder ist berufstätig. Auffällig ist, dass die meisten Nutzer Abitur oder einen Hochschulabschluss als höchsten Bildungsabschluss angeben. Der Zweck der Fernbusfahrt wird meist mit Freizeitaktivitäten beschrieben, wie ein Besuch bei Freunden oder Verwandten. Der Nutzer fährt mehrmals im Jahr mit dem Fernbus und ist ÖV-affin. Nur wenige besitzen einen eigenen PKW. Die optimale Lage einer Fernbushaltestelle wird von den Kunden in der Innenstadt oder in Bahnhofsnähe mit sehr guter Erreichbarkeit mit dem ÖPNV beschrieben. Die Autobahnnähe ist ein eher unwichtiges Kriterium. Ebenfalls sind Haltestellen am Stadtrand nicht gern gesehen. Eine Ausnahme besteht dann, wenn eine gute ÖV-Anbindung vorhanden ist. Die maximale Reisezeit zur Haltestelle liegt bei 50 Minuten, der Durchschnitt sind 38 Minuten Reisezeit. Des Weiteren kann festgestellt werden, dass eine Verknüpfung von Fernbus und ÖV-Ticket von dem Großteil der Befragten auch für einen Aufpreis von zwei oder mehr Euro angenommen werden würde. (Selzer (2016), S. 28-41).

In dem Vortrag „Planung, Finanzierung und Betrieb von Fernbusterminals in Deutschland“ von C. Bange aus dem Jahr 2013, welcher bei der Konferenz „Kommunales Infrastruktur-Management“ gehalten wurde, werden ebenfalls optimale Ausstattung und Lage von Fernbusterminals diskutiert. Dabei werden sowohl Kunden- und Unternehmenssicht als auch die Meinung der Städte vorgestellt. So wird zunächst die Frage nach der Notwendigkeit für Terminals für die verschiedenen Gruppen betrachtet. Für die Busunternehmen stellen Terminals eine garantierte Haltestelle dar, welche auch den Service, wie Werkstatt oder Parkplätze, beinhaltet. Kunden möchten vor allem Terminals, die Ticketschalter, beheizte Wartehallen und dynamische Reiseinformationen benutzen. Diese sind sehr kundenfreundlich. Ein Fernbusterminal ist für die Stadt sinnvoll, da unkoordinierte Haltestellensituationen vermieden werden. Des Weiteren fördern sie Tourismus und Erreichbarkeit der Stadt. So liegt das Interesse der Stadt vor allem im verkehrsplanerischen Sektor. (Bange (2013), S. 3) Wird der Standort einer Haltestelle betrachtet kann ebenfalls zwischen zentral und dezentral unterschieden werden, wobei die Vor- und Nachteile dieselben sind wie bei Stelzer (2016). (Bange (2013), S. 6) Die Fernbusunternehmen bevorzugen eine Haltestelle in der Innenstadt, da dort eine gute ÖV-Anbindung gegeben ist. Des Weiteren gehen die Unternehmen davon aus, dass die Fahrgäste mit Zeitverlusten durch Fahrten in die Innenstädte rechnen und daher dieser Punkt nicht ausschlaggebend ist. (Bange (2013), S. 7) Für Städte und Betreiber der Terminals

kommen sowohl zentrale als auch dezentrale Standorte in Frage. Wichtig hierbei ist eine gute ÖV-Anbindung (Bange (2013), S. 8).

In einer weiteren Studie, die von der FlixBus GmbH zusammen mit der deutschen Hochschule in Heilbronn im Jahr 2014 durchgeführt wurde, geht es um die Kundenzufriedenheit bezüglich der deutschen Fernbusbahnhöfe. Dabei wurden in zwei Wochen 5.531 Personen zu Standort- und Ausstattungsmerkmalen und dem bevorzugten Verkehrsmittel zur An- und Abreise befragt. So kann festgestellt werden, dass mehr als die Hälfte der Befragten den öffentlichen Verkehr zur An- und Abreise zur Fernbushaltestelle nutzen. Weitere Ergebnisse können in der nachfolgenden Tabelle eingesehen werden:

Verkehrsmittel	Anteil
Öffentliche Verkehrsmittel	56,7 %
Zug	12,7 %
Taxi	1,1 %
Fuß	6,2 %
Fahrrad	1,0 %
PKW (gebracht)	20,2 %
PKW (selbst gefahren)	1,5 %
Sonstige	0,5 %

Tabelle 2: Anteile der verwendeten Verkehrsmittel zur An- und Abreise zur Haltestelle (e.D. nach FlixBus (2014))

Es zeigt sich, dass nur wenige zu Fuß oder mit dem Fahrrad anreisen. Dies kann an der Größe des Gepäcks liegen. Ebenfalls kann festgestellt werden, dass das zweithäufigste benutzte Verkehrsmittel der PKW ist, allerdings als Mitfahrer und nicht als Fahrer. Des Weiteren wird nach wichtigen Standortmerkmalen von Fernbushaltestellen gefragt. Die Ergebnisse sind in Tabelle 3 abgebildet.

Merkmale	Anteil
ÖPNV-Anbindung	49,3 %
Nähe zum Bahnhof	37,7 %
Nähe zum Stadtzentrum	9,0 %
Nähe zur Autobahn	1,2 %
Ausreichend Parkplätze	2,9 %

Tabelle 3: Wichtigkeit der Standortmerkmale von Fernbushaltestellen (e.D. nach FlixBus (2014))

Es zeigt sich, dass eine Anbindung zum öffentlichen Nahverkehr für Kunden das wichtigste Merkmal darstellt. An zweiter Stelle steht die Nähe zu einem Bahnhof. Unwichtig ist dabei, ob der Standort in der Innenstadt liegt oder an der Autobahn. Auch ausreichend Parkplätze stellen kein wichtiges Merkmal dar, da die An- und Abreise nur sehr selten mit dem eigenen PKW erfolgt. (FlixBus (2014))

Fitzner und Schmidt (2016) beschäftigen sich mit der Bestimmung von kaufrelevanten Produkteigenschaften von Fernbussen. In der Untersuchung wird eine Conjoint-Analyse verwendet. Dabei wird ebenfalls die Zielgruppe von Fernbussen untersucht. Es kann festgestellt werden, dass Fernbusnutzer preisbewusste Alleinreisende unter 35 Jahren sind. Der Fahrtzweck ist meist von privater Natur. Zwischen den Fernbusunternehmen herrscht starker Wettbewerb. Einzelne kleinere Anbieter sind bereits vom Markt verdrängt worden. Die Unternehmen bleiben nur wettbewerbsfähig, wenn sie Preise und Kosten minimal halten. Mittels Conjoint-Analyse kann der Preis als wichtigstes Merkmal festgestellt werden. Andere Merkmale, die die Kaufentscheidung beeinflussen, sind Streckenangebot, Fahrzeit, Frequenz, Kundenservice und Komfort. (Fitzner (2016), S. 1-26)

3. Methodisches Vorgehen

Im dritten Kapitel dieser Arbeit wird der verwendete Fragebogen erläutert und die daraus resultierenden Daten spezifiziert. Des Weiteren werden die Methoden, welche in Kapitel 4 angewendet werden, vorgestellt.

3.1 *Datenspezifikation*

Zu Beginn soll der in der Umfrage verwendete Fragebogen vorgestellt und näher erläutert werden. Daraufhin werden die daraus resultierenden Daten, die in Kapitel 4 untersucht werden, spezifiziert.

3.1.1 Fragebogen

Der Fragebogen gliedert sich in zwei Teile. Im ersten Teil werden die soziodemografischen und sozioökonomischen Merkmale einer Person bestimmt, die in Kapitel 4.1 für eine Zielgruppenanalyse verwendet werden. Dabei sind folgende Variablen berücksichtigt worden:

- Geschlecht (*binär*)
- Alter in Altersklassen (*ordinal*)
- Beruf (*nominal*)
- Monatliches Nettoeinkommen in Einkommensklassen (*ordinal*)
- Höchster Schulabschluss (*nominal*)
- Höchste Berufsausbildung (*nominal*)
- Verfügbarkeit PKW (*binär*)
- Fernbusnutzer (*binär*)
- Fernbusnutzung, wenn Fernbusnutzer (*ordinal*)

Die verschiedenen Ausprägungen aller Variablen werden in Kapitel 4.1 genauer erläutert und sind ebenfalls im Fragebogen, welcher in Anhang A.1 zu finden ist, dargestellt. Im zweiten Teil des Fragebogens werden die Auswahlentscheidungen, welche in Kapitel 4.2 für die auswahlbasierte Conjoint-Analyse benötigt werden, beantwortet. Hierbei ist zu beachten, dass jeweils vier lange Strecke (mit mehreren Zwischenhalten) und vier kurze Strecke (ohne Zwischenhalt) betrachtet werden. Für jede Strecke stehen zwei Varianten zur Verfügung, zwi-

schen denen die Probanden wählen können. Eine „Non-Option“ ist nicht vorhanden. Des Weiteren ist zu beachten, dass die Probanden keine ÖV-Tickets wie Semestertickets, Jobtickets oder Monatsfahrkarten zur Verfügung stehen haben, sodass sie sich für die Fahrten vom Startpunkt zum Fernbusbahnhof und vom Fernbusbahnhof zum Endpunkt zusätzliche Tickets kaufen müssen. Ausnahme bilden hier spezifisch in den Varianten angegebene Fahrkarten, welche nur auf der Strecke vom Fernbusbahnhof zum Endpunkt gültig sind. Für die Varianten sind folgende Informationen angegeben:

- Reisezeit im Bus
- Kosten
- Reisezeit Startpunkt bis Haltestelle
- Reisezeit Haltestelle bis Zielpunkt
- Komplexe Reisezeit
- Verfügbarkeit extra ÖV-Ticket für die Strecke Haltestelle bis Zielpunkt

Dabei beschreibt die Variante 1 jeweils die Situation mit Fernbusbahnhöfen innerhalb der Städte und Variante 2 die Situation, dass die Fernbusbahnhöfe in der Nähe von Autobahnen an Haltestellen des öffentlichen Nahverkehrs liegen. Um die Einsparung der Zeit für Variante 2 deutlich zu machen, werden pro Zwischenstopp 30 Minuten von der Reisezeit im Bus abgezogen. Die Reisezeit für Variante 1 ist jeweils die längste für die betrachtete Strecke ohne Umstieg. Allerdings ist der Preis der Variante 2 jeweils um drei Euro höher als in Variante 1, welcher der durchschnittliche Preis für die jeweilige Strecke ist.¹ Für die Reisezeiten von und zu den Haltestellen werden für jede Variante andere Zeiten berücksichtigt, um die unterschiedlichen Standorte der Fernbushaltestellen darzustellen. Dabei werden für beide Varianten keine Wartezeiten mit einberechnet, um die direkte Anbindung an den ÖPNV zu simulieren. Die Zwischenstopps der einzelnen Strecken sind in Anhang A.2 zu finden. Ebenfalls ist zu beachten, dass sich die Reisezeiten von Startpunkt bis Haltestelle und Haltestelle bis Zielort unterscheiden, um die Änderung der Städte zu berücksichtigen.

¹ MeinFernbus, Die Reisezeiten und Kosten für die jeweiligen Strecken sind mit Hilfe von der Internetseite von MeinFernbus Flixbus (www.meinfernbus.de) recherchiert worden. Der Stand der Daten, welcher im Fragebogen verwendet wurde, ist der 24.03.2017 und kann sich durchaus zum heutigen Stand unterscheiden.

3.1.2 Daten

Die Daten, die hier untersucht werden sollen, sind mittels in Kapitel 3.1.1 vorgestellten Fragebogens erhoben worden. Insgesamt haben 129 Personen den Fragebogen beantwortet. Davon sind zehn Fragebögen unvollständig und somit beläuft sich die Antwortquote auf 92,25 Prozent. Der Fragebogen war vom 25.03.2017 bis zum 23.04.2017 online und somit frei zugänglich. Die Probanden sind mittels sozialer Netzwerke auf den Fragebogen aufmerksam gemacht worden, sodass eine räumliche Eingrenzung schwer möglich ist. Die Stichprobe beläuft sich also auf $n = 119$. Für die Zielgruppenanalyse werden alle soziodemographischen Merkmale verwendet. Dabei werden die Probanden in Fernbusnutzer $n_{FBN} = 93$ und in Nicht-Nutzer $n_{NN} = 26$ aufgeteilt. In der Conjoint-Analyse sollen dann die Entscheidungen der Probanden untersucht werden, sodass alle im Fragebogen enthaltenden Daten verwendet werden. Zusätzlich wird in Kapitel 4 der Einfluss der soziodemografischen Merkmale auf die Auswahlentscheidungen analysiert.

3.2 Methoden

Im Folgenden sollen grundlegende Annahmen der Conjoint-Analyse und der logistischen Regression vorgestellt werden. Im Anschluss werden die in Kapitel 4 verwendete auswahlbasierte Conjoint-Analyse und die logistische Regression näher erläutert.

3.2.1 Allgemein

In dieser Arbeit sollen die mittels Fragebogen erhobenen Daten zunächst mit einer Zielgruppenanalyse untersucht werden. Hierbei werden verschiedene soziodemografische Merkmale wie z.B. Geschlecht, Alter, Einkommen oder Beruf betrachtet. Es soll ein Unterschied zwischen den Personen, die Fernbusse nutzen (Nutzer), und Personen, die Fernbusse nicht nutzen (Nichtnutzer), innerhalb dieser Variablen festgestellt werden, um möglichst eine genaue Definition einer Zielgruppe für Fernbusunternehmen festlegen zu können. Dabei werden einfache deskriptive Größen und Verteilungen in den einzelnen Gruppen verwendet. Im Anschluss daran wird mittels auswahlbasierter Conjoint-Analyse die Akzeptanz von Fernbusbahnhöfen in Routennähe untersucht. Conjoint-Analysen können als Individualanalysen definiert werden, welche die Beurteilung einer einzelnen Person zu einer großen Menge an Objekten abbildet. Dabei kann die Beurteilung oder Auswahlentscheidung einer Person für verschiedene Objekte,

welche mittels mehrerer Merkmale beschrieben werden, festgestellt werden (Backhaus (2016), S. 518). Im Allgemeinen wird zwischen traditionellen und auswahlbasierten Conjoint-Analysen unterschieden. Während bei der traditionellen Analyse die zu untersuchenden Objekte in eine Reihenfolge gebracht werden müssen, welche die Präferenzen der Person widerspiegelt, ist das Ziel der auswahlbasierten Analyse eine Auswahl aus verschiedenen Alternativen (Backhaus (2016), S. 518). Mit Hilfe der Analysen kann ein Gesamtnutzenwert für jedes Objekt bestimmt werden, welcher sich aus Teilnutzenwerten für jede Ausprägung berechnen lässt. Die Conjoint-Analyse gehört zu den dekompositionellen Verfahren der Präferenzmessung. Dies bedeutet, dass sich die Entscheidung einer Person auf ein ganzheitliches Objekt bezieht (CONsidered JOINTly), daraus aber die Teilnutzenwerte für jede Ausprägung bestimmt werden können. Neben den dekompositionellen existieren noch die kompositionellen und die hybriden Verfahren, welche zur Präferenzmessung genutzt werden können. In Abbildung 1 sind die verschiedenen Untergruppen dieser drei Verfahren nach Sattler (1999) dargestellt (Backhaus (2016), S. 518/519, Fabian(2005), S. 116/117).

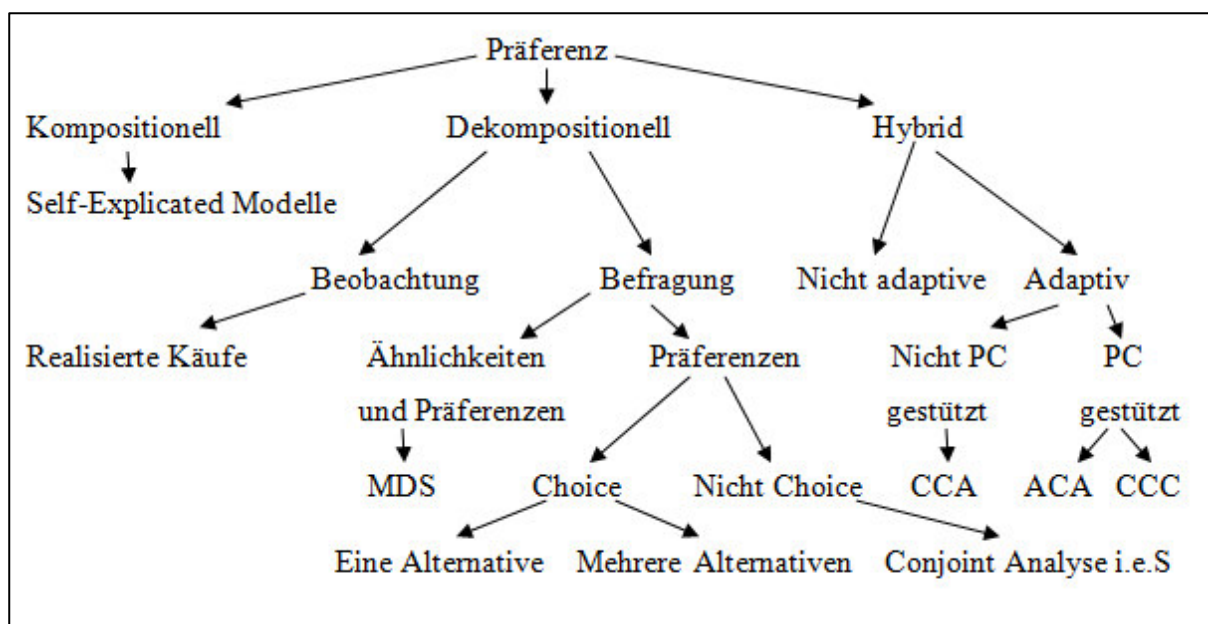


Abbildung 1: Methoden zur Präferenzmessung (Fabian (2005), S. 116)

In dieser Arbeit soll die auswahlbasierte Conjoint-Analyse verwendet werden, welche im folgenden Kapitel 3.2.2 genau erläutert wird.

Das zweite in Kapitel 4 verwendete Verfahren ist die logistische Regression. Diese gehört zu den struktur-prüfenden Verfahren und ist eine Variante der Regressionsanalyse mit dem Un-

terschied, dass die abhängige Variable binär kodiert ist. Die endogene Variable stellt meist zwei Ereignisse dar und wird als Zufallsvariable betrachtet. Dies wird als binäre logistische Regression bezeichnet. Ziel ist es, für beide Ereignisse die Eintrittswahrscheinlichkeit vorherzusagen, für die gilt:

$$\text{prob}(Y = 0) = 1 - \text{prob}(Y = 1) \quad (\text{Backhaus (2011), S. 250/251}) \quad (3.1)$$

Die exogenen Variablen werden mittels Linearkombination miteinander verknüpft, welche als systematische Komponente bezeichnet wird:

$$z_g = \beta_0 + \sum_{j=1}^J \beta_j * x_{jg} + o_g \quad (3.2)$$

Dabei ist z der Logitwert für jede Auswahlentscheidung g . Die Variable j bezeichnet die Eigenschaften des Produktes und o steht für den Fehlerterm. So ist y_g dann gleich 1, wenn $z_g > 0$, sonst nicht. Dabei wird die Wahrscheinlichkeit (*prob*) mit einer logistischen Funktion berechnet:

$$\text{prob} = \frac{e^z}{1+e^z} = \frac{1}{1+e^{-z}} \quad \text{mit } e = 2,7182 \text{ (Eulersche Zahl)} \quad (3.3)$$

Die logistische Funktion besitzt einen S-förmigen Verlauf mit dem Schnittpunkt der y-Achse bei 0,5. Sie kann als Verteilungsfunktion interpretiert werden. (Backhaus (2016), S. 284/285)
Die genaue Vorgehensweise wird in Kapitel 3.2.3 erklärt.

3.2.2 Auswahlbasierte Conjoint-Analyse

Bei einer auswahlbasierten Conjoint-Analyse werden die Präferenzen von Personen bei der Beurteilung von Stimuli indirekt durch Auswahlentscheidungen ermittelt. Das Ziel liegt darin, die Relevanz von Eigenschaften eines Produktes zu ermitteln. Bei diesem Verfahren stehen realitätsnahe Varianten zur Auswahl, welche durch die nominale Skalierung zu Informationsverlusten führt (Backhaus (2015), S. 176/177). Die auswahlbasierte Conjoint-Analyse oder auch Choice Based Conjoint-Analyse (CBCA) gliedert sich grob in die Teile Erhebungsdesign und Analyseverfahren. In der Phase des Erhebungsdesigns werden grundsätzliche Vorgänge der Datenermittlung geklärt. Dies bedeutet, dass zum einen Umfang und Art der Stichprobengewinnung bestimmt und zum anderen die zu verwendenden Stimuli und Choice Sets gestaltet werden. Anhand dieser Überlegungen können dann Fragebögen oder andere Maßnahmen für die Ermittlung der Daten erstellt werden. Im Analyseverfahren werden sowohl

verhaltenstheoretische Modelle spezifiziert als auch entschieden, mit welchen statistischen und mathematischen Verfahren die Schätzung des Modells durchgeführt wird (Backhaus (2015), S. 180). Im Allgemeinen kann die CBCA in folgende Schritte unterteilt werden:

1. Gestaltung der Stimuli
2. Gestaltung der Auswahlsituation
3. Spezifikation des Nutzenmodells
4. Spezifikation eines Auswahlmodells
5. Schätzung der Nutzenwerte
6. Interpretation und Anwendung
7. Disaggregation der Nutzenwerte

Dabei zählen die Schritte eins und zwei zum Erhebungsdesign und die Schritte drei, vier und fünf zum Analyseverfahren (Backhaus (2015), S. 180/181). Die einzelnen Schritte sollen im Folgenden näher erläutert werden.

Gestaltung der Stimuli

Hier wird die Anzahl der Eigenschaften eines Produktes und deren Ausprägungen ermittelt. Neben den Alternativen für die vorhandenen Produkte empfiehlt sich eine „Non-Option“ in jedem Choice Set. Fällt diese weg, wird dies „forced choice“ genannt. Dies bedeutet, dass der Proband zu einer Wahl gezwungen ist, auch wenn er mit der Auswahl der Alternativen nicht einverstanden ist. Die verschiedenen Stimuli können verbal, visuell oder physisch dem Probanden vorgestellt werden. Die Auswahl der Methode hängt von der Art der Befragung ab. Es sollte darauf geachtet werden, dass nicht zu viele Stimuli verwendet werden. Mit steigender Zahl der Eigenschaften und deren Ausprägung steigt die Anzahl der Stimuli exponentiell an, sodass der Aufwand der Befragung ansteigt und auch die Auswahl für den Probanden schwieriger wird. (Backhaus (2015), S.181/182).

Gestaltung der Auswahlsituation

„Eine Auswahlsituation ist eine Menge an Stimuli (Alternativen), zwischen denen eine Testperson eine Auswahl trifft, also entscheiden soll. Sie besteht aus dem Choice Set und der Fragestellung“ (Backhaus (2015), S. 183). Während bei der traditionellen Conjoint-Analyse dem Probanden alle Stimuli zur gleichen Zeit vorgelegt werden, bekommt er bei der CBCA nur eine Teilmenge zu sehen. Nach der Entscheidung werden ihm solange weitere Teilmengen

vorgestellt, bis es keine mehr gibt. Dabei muss auf die Größe und die Anzahl der Choice Sets geachtet werden (Backhaus (2015), S. 183). Generell werden in diesem Schritt folgende Festlegungen getroffen:

- Anzahl der Alternativen je Choice Set
- Anzahl der Choice Sets je Testperson
- Verwendung der Non-Option
- Zuordnung von Choice Sets zu Testpersonen

Nach der Festlegung der Auswahlssituation kann diese den Probanden vorgelegt werden. Die getroffenen Entscheidungen können in nachfolgenden Analysen so untersucht werden, dass die Teilnutzenwerte der einzelnen Ausprägungen festgestellt werden können. Daraus werden die einzelnen Präferenzen ersichtlich. (Backhaus (2015), S. 185/186)

Spezifikation des Nutzenmodells

Das Nutzenmodell wird dafür verwendet, die Entscheidung der Personen nachvollziehen zu können. Dabei sollen die Präferenzen der jeweiligen Person ermittelt werden. Es existieren drei Teilnutzenmodelle, welche den Zusammenhang von Ausprägung und Nutzen erklären können: das Vektor-Modell, das Idealpunkt-Modell und das Teilwert-Modell. Im Vektor-Modell wird ein linearer Zusammenhang von Ausprägung und Nutzen angenommen. Dies bedeutet, dass je mehr von der Ausprägung vorhanden ist, desto größer der Nutzen für die Person. Das Idealpunkt-Modell geht von einem optimalen Punkt aus, der für jede Ausprägung erreicht werden kann. Zu wenig oder zu viel mindert den Nutzen wieder. Das Teilwert-Modell ist diskret, daher kann es zu Informationsverlust kommen. Allerdings können andere Nutzenmodelle approximiert werden und somit ist es sehr flexibel. (Backhaus (2015), S. 187) Aus verschiedenen Teilnutzenmodellen kann mittels additiver oder multiplikativer Verknüpfung ein Gesamtnutzen-Modell erstellt werden. Das Gängigste ist das additive Teilwert-Nutzenmodell, welches wie folgt dargestellt werden kann:

$$u_{st} = \sum_{j=1}^J \sum_{m=1}^{M_j} b_{jm} * w_{jms} \quad (3.4)$$

Mit u_{st} = Nutzen von Stimuli st , b_{jm} = Teilnutzen einer Ausprägung m mit Eigenschaft j und w_{jms} = Dummy-Variable ($w_{jms} = 1$, falls Stimulus s bezüglich Eigenschaft j die Ausprägung m hat, sonst 0) (Backhaus (2015), S. 189).

Ein additives Nutzenmodell wird auch kompensatorisch genannt, bei dem eine Eigenschaft, welche einen geringeren Nutzenbeitrag besitzt, von einer anderen mit einem hohen Nutzenbeitrag kompensiert wird. Wird ein multiplikatives Modell verwendet, wird der Gesamtnutzen Null, wenn eine Eigenschaft einen Teilnutzen von Null besitzt. Im Verfahren der auswahlbasierten Conjoint-Analyse werden mehrere Auswahl-situationen betrachtet, sodass sich folgende Nutzenfunktion ergibt:

$$u_{kr} = \sum_{j=1}^J \sum_{m=1}^{M_j} b_{jm} * w_{jmkr} \quad (3.5)$$

Mit u_{kr} = Nutzen der Alternative k in Auswahl-situation r , b_{jm} = Teilnutzen einer Ausprägung m mit Eigenschaft j und $w_{jmkr} = 1$, falls Alternative k in Situation r bezüglich Eigenschaft j die Ausprägung m hat, sonst 0. (Backhaus (2015), S. 189)

Spezifikation eines Auswahlmodells

Ergänzend zu dem Nutzenmodell wird für eine CBCA ein verhaltenstheoretisches Modell, wie ein Auswahlmodell, benötigt. Dies ist notwendig, da hier eine Entscheidung zwischen Alternativen getätigt wird und nur mit Hilfe eines weiteren Modells erklärt werden kann, „...wie sich eine Person auf Basis ihrer Nutzenvorstellungen bei der Auswahl zwischen Alternativen entscheidet.“ (Backhaus (2015), S. 191). Dabei kann zwischen mehreren Auswahlmodellen, wie z.B. dem Max-Utility-, dem Random-Choice-, dem Attraction- oder dem Logit-Choice-Modell, gewählt werden. Choice-Modelle vereinfachen stark das menschliche Entscheidungsverhalten, sodass keine exakten Entscheidungen, sondern nur Wahrscheinlichkeiten für die Wahl einer Alternative abgebildet werden können. Im Max-Utility-Modell wird der Alternative mit dem höchsten Nutzen die Wahrscheinlichkeit 1 zugeordnet, während im Random-Choice-Modell alle Alternativen die gleiche Wahrscheinlichkeit zugeordnet bekommen. Das Attraction-Modell ordnet die Auswahlwahrscheinlichkeit proportional den Nutzenwerten zu. Das Logit-Choice-Modell ist das bekannteste und findet die häufigste Anwendung. Dabei werden die Wahrscheinlichkeiten für jede Alternative durch die Differenzen der Nutzenwerte bestimmt:

$$prob_i(k|k' \in CS) = \frac{e^{\rho_i * u_{ik}}}{\sum_{k \in CS} e^{\rho_i * u_{ik}}} = \frac{1}{1 + \sum_{k' \neq k \in CS} e^{-\rho_i * (u_{ik} - u_{ik'})}} \quad (3.6)$$

CS steht hier für die Choice Sets und ρ ist der Rationalitätsparameter (Backhaus (2015), S. 191/192)

Im Logit-Choice-Modell liegen die Auswahlwahrscheinlichkeiten zwischen 0 und 1, und können daher mit einer logistischen Funktion dargestellt werden. Stehen nur zwei Alternativen zur Verfügung handelt es sich um ein binäres Logit-Choice-Modell. Hier wird die Wahrscheinlichkeit für die Wahl von Alternative 1 gegenüber Alternative 2 mit Hilfe der Nutzenwerte ermittelt:

$$prob(1|2) = \frac{e^{u_1}}{e^{u_1} + e^{u_2}} = \frac{1}{1 + e^{-(u_1 - u_2)}} \quad (3.7)$$

Und damit

$$prob(2|1) = 1 - prob(1|2) \quad (3.8)$$

Daraus kann nun der *Logit* z mit Hilfe des Chancen-Verhältnisses (*Odds*) berechnet werden:

$$Logit = \ln\left(\frac{prob}{1-prob}\right) \quad (3.9)$$

(Backhaus (2015), S.193)

Die charakteristischen Eigenschaften eines Logit-Choice-Modells lassen sich wie folgt zusammenfassen:

- Die Auswahlwahrscheinlichkeit für eine Alternative ist abhängig von ihrem und allen anderen Nutzen.
- Es zählt die Nutzendifferenz und nicht der absolute Nutzen.
- Durch kleine Änderungen der Nutzenwerte können sich die Auswahlwahrscheinlichkeiten bei ähnlichen Alternativen stark verändern. Sind die Nutzendifferenzen sehr groß, haben Änderungen nur eine kleine Auswirkung auf die Wahrscheinlichkeiten.
- Das Wahrscheinlichkeitsverhältnis ist unabhängig von der Anzahl der Alternativen.

(Backhaus (2015), S. 194)

Schätzung der Nutzenwerte

Um die Teilnutzenwerte b_{jm} zu schätzen, wird die Maximum-Likelihood-Methode verwendet. Da die abhängige Variable nicht beobachtbar ist und nur die Auswahlentscheidungen als nominale Größen vorliegen, kann keine Regressionsanalyse durchgeführt werden. Die Likelihood-Funktion kann wie folgt dargestellt werden:

$$L = \prod_{r=1}^R \prod_{k=1}^K prob_r(k)^{d_{kr}} \Rightarrow \underset{b_j}{Max!} \quad (3.10)$$

Wobei $d_{rk} = 1$, wenn in Situation r die Alternative k gewählt wurde, sonst $d_{rk} = 0$. Durch Logarithmieren kann diese Funktion vereinfacht werden, sodass folgende LogLikelihood-Funktion entsteht:

$$LL = \sum_{r=1}^R \sum_{k=1}^K \ln(prob_r(k)) * d_{rk} \Rightarrow \underset{b_j}{Max!} \quad (3.11)$$

Mit

$$prob_r(k) = \frac{e^{u_{kr}}}{\sum_{k' \in CS_r} e^{u_{k'r}}} \quad (\text{Choice-Modell}) \quad (3.12)$$

$$u_{kr} = \sum_{j=1}^J \sum_{m=1}^{M_j} b_{jm} * w_{jmk_r} \quad (\text{Nutzenmodell}) \quad (3.13)$$

Die Teilnutzen müssen nun so bestimmt werden, dass der LogLikelihood-Wert maximiert wird. Da LL nur negative Werte annehmen kann, ist das Ziel dem Wert 0 so nah wie möglich zu kommen. Dies geschieht mittels Quasi-Newton-Verfahren oder Gradientenverfahren, welche einen iterativen Algorithmus darstellen. (Backhaus (2015), S. 196-198)

Die rechnerische Durchführung der CBCA gliedert sich in folgende Schritte:

1. Ermittlung von Startwerten
2. Verankerung der Teilnutzenwerte
3. Berechnung der Gesamtnutzenwerte
4. Berechnung der Auswahlwahrscheinlichkeiten
5. Maximum-Likelihood-Schätzung
6. Berechnung weiterer Ergebnisse
7. Güteprüfung

Bei der Verwendung von iterativen Algorithmen, werden Startwerte benötigt. Eine Möglichkeit, diese zu berechnen, ist die Bestimmung der absoluten gewählten Häufigkeiten jeder Eigenschaftsausprägung, welche im Anschluss durch die Anzahl der Auswahl-situation T geteilt wird. Daraufhin werden die Teilnutzenwerte verankert. Dabei gibt es zum einen die Möglichkeit eine „Null-Kategorie“ für jede Eigenschaft zu wählen oder die Teilnutzen zu zentrieren, sodass die Summe Null ergibt. Für M Teilnutzen müssen nur $M-1$ Teilnutzen geschätzt werden. (Backhaus (2015), S. 200/201) Daraufhin folgt die Berechnung der Auswahlwahrschein-

lichkeiten mit Formel (3.9). Um nun optimale Schätzwerte zu erhalten, wird mit Hilfe von Formel (3.8) der LogLikelihood-Wert maximiert. Dabei ändern sich die Startwerte b_{jm} . So kann nun der Gesamtnutzen für alle Stimuli berechnet werden. Um die Wichtigkeit einer Eigenschaft zu ermitteln, wird die Spannweite ermittelt. Diese ist unabhängig von Nullpunktverschiebungen und kann durch Normierung in relative Wichtigkeiten gewandelt werden, sodass Ergebnisse von unterschiedlichen Analysen verglichen werden können. Zum Schluss wird nun die Güte des Modells geprüft. Diese kann mittels Likelihood-Ratio-Statistik (LLR) oder mit Hilfe des McFadden R^2 (R^2_M) ermittelt werden:

$$LLR = -2 * \ln\left(\frac{L_0}{L_V}\right) = 2 * (LL_0 - LL_V) \quad (3.14)$$

Mit L_0 = Likelihoodwert des Nullmodells, L_V = Likelihoodwert des vollständigen Modells, LL_0 = LogLikelihoodwert des Nullmodells und LL_V = LogLikelihoodwert des vollständigen Modells.

$$R^2_M = 1 - \left(\frac{LL_V}{LL_0}\right) \quad (3.15)$$

Der Likelihood-Ratio Test kann mit dem F-Test der normalen Regressionsanalyse verglichen werden. Im Gegensatz dazu ist der McFadden- R^2 normiert und kann nur Werte zwischen 0 und 1 annehmen. Dieser Wert lässt sich aber nicht als Anteil der Streuung erklären und ist somit ein Pseudo- R^2 . Des Weiteren kann die Trefferquote eine Aussage über die Güte eines Modells geben. Hier wird berechnet, wie viele Auswahl-situationen richtig zugeordnet werden. Anschließend können die geschätzten Parameter auf ihre Signifikanz mit der Nullhypothese $H_0: b_{jm} = 0$ untersucht werden. Es wird ebenfalls der Likelihood-Ratio-Test verwendet. Dabei wird jeweils der zu untersuchende Parameter aus dem Modell ausgeschlossen und ein neuer LL_{0j} berechnet. (Backhaus (2015), S. 204 -211)

Interpretation und Anwendung

Mit Abschluss der Schätzung der Nutzwerte sind alle Schritte des Erhebungsdesigns und des Analyseverfahrens abgeschlossen. Die ermittelten Ergebnisse können nun interpretiert werden. Da nur die Differenzen der Teilnutzenwerte für die Auswahlwahrscheinlichkeiten wichtig sind, werden diese oft als Marktanteile interpretiert. Mit Hilfe dieser Ergebnisse kann ein Unternehmer vorhersagen, wie sich die Nachfrage für Produkte entwickeln kann und entscheiden, in welches Produkt er investieren möchte. (Backhaus (2015), S. 212/213)

Disaggregation der Nutzenwerte

Im Gegensatz zur traditionellen Conjoint-Analyse werden bei der CBCA keine individuellen Entscheidungen betrachtet. Die betrachtete Gruppe wird in der Analyse aggregiert. Eine aggregierte Nutzenfunktion kann aber nur dann verwendet werden, wenn davon ausgegangen werden kann, dass die Probanden weitestgehend die gleichen Vorstellungen von Nutzen haben und diese somit homogen sind. Um Heterogenität zu berücksichtigen, kann unter verschiedenen Verfahren gewählt werden. Zum einen ist eine a priori Segmentierung anwendbar, die die Häufigkeiten von Merkmalsausprägungen verwendet, zum anderen können auch der Latent Class (LC)-Ansatz oder der Hierarchical Bayes (HB)-Ansatz für die Einbeziehung der Heterogenität in Betracht gezogen werden. Im LC-Ansatz wird davon ausgegangen, dass die Stichprobe nicht direkt beobachtbare Gruppen umfasst. Die Probanden können zwar den verschiedenen Gruppen zugeordnet werden, jedoch geschieht dies nur mit einer gewissen Wahrscheinlichkeit. Der HB-Ansatz bedient sich ebenfalls dem Grundsatz bedingter Wahrscheinlichkeiten. Anders als beim LC-Ansatz wird hier eine kontinuierliche Verteilung der Präferenzen angenommen, sodass die verschiedenen betrachteten Gruppen nicht vollständig voneinander getrennt werden können. (Backhaus (2015), S. 215-218)

3.2.3 Logistische Regression

Anschließend wird das Verfahren der logistischen Regression näher betrachtet, damit diese in Kapitel 4 angewendet werden kann, um die Einflüsse der sozioökonomischen Variablen auf die Auswahlentscheidungen zu untersuchen. Die logistische Regression kann in folgende Schritte aufgeteilt werden:

- Modellformulierung
- Schätzung der logistischen Regressionsfunktion
- Interpretation der Regressionskoeffizienten
- Prüfung des Gesamtmodells
- Prüfung der Merkmalsvariablen

Diese Punkte werden im Folgenden einzeln betrachtet und erklärt. (Backhaus (2011), S. 257)

Modellformulierung

Zuerst müssen die möglichen Ereignisse der abhängigen Variablen und die unabhängigen Einflüsse festgelegt werden. Anschließend wird das Regressionsmodell ausgesucht, welches für die Analyse verwendet werden soll. Des Weiteren können verschiedene Hypothesen aufgestellt werden, die den Zusammenhang zwischen den endogenen und exogenen Variablen beschreiben. Dabei ist darauf zu achten, dass diese so formuliert werden, dass der Zusammenhang von unabhängigen Variablen und Eintrittswahrscheinlichkeiten erklärt wird. Des Weiteren zeigt sich eine nicht-linearere Beziehung, welche durch die s-förmige Wahrscheinlichkeitsfunktion erklärt werden kann. (Backhaus (2011), S. 257)

Schätzung der logistischen Regressionsfunktion

Für die logistische Regression wird wie auch bei der Conjoint-Analyse die Maximum Likelihood-Schätzung verwendet, welche die Parameter a_{eig} so schätzt, dass die Wahrscheinlichkeiten maximal werden. Die Wahrscheinlichkeiten einer betrachteten Auswahlwahrscheinlichkeit g sollten dabei entweder $prob(Y=1)$ oder $prob(Y=0)$ sein. Dies ist genau dann der Fall, wenn

$$prob_g(y) = \left(\frac{1}{1+e^{-z_g}} \right)^{y_g} * \left(1 - \frac{1}{1+e^{-z_g}} \right)^{1-y_g} \quad (3.16)$$

Dabei ist der erste Teil der Terms gleich 1, wenn $y_g = 0$ und somit ergibt sich eine Wahrscheinlichkeit von $prob_g(y=0)$. Wenn $y_g = 1$, dann ist der zweite Teil des Terms gleich 1 und somit ist $prob_g(y=1)$. Die Wahrscheinlichkeiten sollen dabei maximal werden. Für die Zuweisung zu einer Kategorie wird der Wert 0,5 verwendet. Ist die Wahrscheinlichkeit größer als 0,5 wird der Fall zu $y = 1$ zugeordnet. Bei einer Wahrscheinlichkeit kleiner als 0,5 geschieht die Zuweisung zu $y = 0$. Mit Hilfe der Likelihood-Funktion sollen die Einzelwahrscheinlichkeiten gleichzeitig maximiert werden:

$$L = \prod_{g=1}^G \left(\frac{1}{1+e^{-z_g}} \right)^{y_g} * \left(1 - \frac{1}{1+e^{-z_g}} \right)^{1-y_g} \xRightarrow{a_{eig}} max \quad (3.17)$$

Dabei kann ein Logarithmus verwendet werden, um das Maximierungsproblem einfacher zu gestalten:

$$LL = \sum_{g=1}^G \left[y_g * \ln \left(\frac{1}{1+e^{-z_g}} \right) \right] + \left[(1 - y_g) * \ln \left(1 - \frac{1}{1+e^{-z_g}} \right) \right] \xRightarrow{a_{eig}} max \quad (3.18)$$

Diese Likelihood-Funktion wird meist mit dem Newton-Raphson-Algorithmus maximiert. (Backhaus (2011), S. 258-260)

Interpretation der Regressionskoeffizienten

Nach der Regressionsschätzung können nun die Koeffizienten interpretiert werden. Da aber kein linearer Zusammenhang zwischen den exogenen Variablen und der Wahrscheinlichkeit besteht, ist dies schwierig. So können weder die Koeffizienten untereinander verglichen werden, noch kann eine genaue Wirkung der unabhängigen Variablen bestimmt werden. Die Parameter haben dennoch einen Einfluss auf die logistische Funktion. Der Parameter a_0 ist für eine Rechts- oder Linksverschiebung der Kurve zuständig, während die Koeffizienten a_{eig} die Steigung bestimmen. Die Werte der Parameter können dementsprechend nicht als globales Maß für die Einflussstärke der unabhängigen Variablen auf die Wahrscheinlichkeiten betrachtet werden. Es ist also nur die Richtung der Einflüsse erkennbar. Um die Koeffizienten dennoch interpretieren zu können, werden die Wahrscheinlichkeitsverhältnisse zwischen Eintrittswahrscheinlichkeit und Gegenwahrscheinlichkeit betrachtet. Diese werden auch *Odds* genannt:

$$Odds(y = 1) = \frac{prob(y=1)}{1-prob(y=1)} = e^z \quad (3.19)$$

Mit Hilfe der *Odds* können die sogenannten *odds ratio* mit $e^{a_{eig}}$ bestimmt werden. Diese werden als Effekt-Koeffizienten bezeichnet und zeigen die Veränderung der *Odds* bei einer Erhöhung einer unabhängigen Variable um eine Einheit. (Backhaus (2011), S. 262-265)

Prüfung des Gesamtmodells

Die Prüfung des Gesamtmodells beschäftigt sich vor allem mit den Fragen, wie gut die Parameterschätzungen das gesamte Modell darstellen und ob Ausreißer vorliegen, die eventuell aus der Analyse ausgeschlossen werden sollten. Um das Gütemaß des Modells zu bestimmen können ein LogLikelihood-Ratio-Test, verschiedene Pseudo-R²-Statistiken wie McFadden-R², Cox and Snell- R² oder Nagelkerke-R² herangezogen werden oder die Klassifikationsergebnisse beurteilt werden. Sowohl der Likelihood-Ratio-Test als auch das McFadden-R² sind in Kapitel 3.2.2 bereits erklärt worden. Der Cox and Snell-R²-Wert kann nur Werte kleiner eins annehmen. Ein Modell kann also nie eine Anpassung von 100 Prozent erlangen. Der

Nagelkerke- R^2 -Wert kann den maximalen Wert von eins erreichen und somit auch eindeutig interpretiert werden.

$$\text{Cox and Snell} - R^2 = 1 - \left(\frac{L_0}{L_V} \right)^{\frac{2}{n-a}} \quad (3.20)$$

$$\text{Nagelkerke} - R^2 = \frac{\text{Cox and Snell} - R^2}{R_{\max}^2} \quad \text{mit } R_{\max}^2 = 1 - (L_0)^{\frac{2}{n-a}} \quad (3.21)$$

Wird die Methode der Klassifizierungsbeurteilung angewendet, werden die empirisch beobachteten Gruppenzuordnungen mit denen der Regressionsgleichung verglichen. Dabei wird als Trennwert wiederum der Wert 0,5 verwendet. So kann die Trefferquote der richtig zugeordneten Personen bestimmt werden. Die Ergebnisse der Klassifizierung können in einer Klassifikationsmatrix dargestellt werden. (Backhaus (2011), S. 267 -273)

Die zweite Frage, welche sich dem gesamten Gütemaß widmet, ist die Ausreißerdiagnostik. Dabei sollen die einzelnen Effekte von Auswahlentscheidungen auf die Gesamtgüte des Modells untersucht werden. Es können zwei Gründe für eine schlechte Anpassung des Modells festgestellt werden. Zum einen kann es sein, dass die exogenen Variablen die endogene nicht beeinflussen und zum anderen kann eine große Anzahl an Beobachtungen vorliegen, die den Modellzusammenhang nicht aufweisen und somit das Ergebnis verzerren. Auskunft über Ausreißer können die individuellen Residuen geben:

$$\text{RESID}_g = y_g - \text{prob}_g(y) \quad (3.22)$$

Die Residuenwerte liegen immer zwischen -1 und 1. Bei Betragswerten von deutlich höher als 0,5 kann laut Literatur davon ausgegangen werden, dass es sich um einen Ausreißer handelt. Mit Hilfe von Gewichten können die standardisierten Residuen bestimmt werden, welche Ausreißer besser erkennbar machen:

$$\text{ZResid}_k = \frac{y_g - \text{prob}(y_g=1)}{\sqrt{\text{prob}(y_g=1) * (1 - \text{prob}(y_g=0))}} \quad (3.23)$$

Jeder Ausreißer muss einzeln analysiert und beurteilt werden. Es kann ein atypisches Antwortverhalten einiger Probanden die Ursache sein oder eine schlechte Spezifikation des Modells. (Backhaus (2011), S. 277/278)

Prüfung der Merkmalsvariablen

Wie auch in der Conjoint-Analyse können neben dem Gütemaß des gesamten Modells die Merkmalsvariablen einzeln untersucht werden. Dabei kann vor allem festgestellt werden, ob zu viele exogene Variablen vorliegen und somit das Modell überspezifiziert ist. Für die Überprüfung wird der Likelihood-Quotienten-Test, eine Erweiterung des Likelihood-Ratio-Tests, angewendet. Dieser ist mit dem Likelihood-Ratio-Test für die Merkmalsüberprüfung in der Conjoint-Analyse identisch. Dabei wird das vollständige Modell mit mehreren reduzierten Modellen verglichen. Für jeden Koeffizienten wird ein reduziertes Modell aufgestellt, bei dem das zu betrachtende Merkmal auf null gesetzt wird. Die Signifikanz der einzelnen Merkmale wird mit der χ^2 -Verteilung überprüft. (Backhaus (2011), S. 279)

4. Analysen

Im Folgenden werden die Daten, die aus dem Fragebogen gewonnen werden konnten, analysiert. Zunächst soll eine Zielgruppenanalyse durchgeführt werden, um einen Unterschied zwischen Nutzern des Fernbusmarktes und den Nichtnutzern feststellen zu können. Hierbei wird sich vor allem auf die soziodemografischen Merkmale der Personen konzentriert. Im Anschluss daran soll mittels auswahlbasierter Conjoint-Analyse, welche in Kapitel 3.2 näher beschrieben wurde, die Akzeptanz von Fernbusbahnhöfen in Routennähe mit einer direkten Anbindung an das Netz des ÖPNV untersucht werden. Zum Schluss wird mittels logistischer Regression die Zusammenhänge zwischen personenspezifischen Merkmalen und Auswahlentscheidungen analysiert.

4.1 Zielgruppenanalyse

Im Folgenden sollen die Probanden, die an der Umfrage teilgenommen haben, näher untersucht werden. Dabei werden die soziodemografischen Variablen wie Geschlecht, Alter, Einkommen oder Beruf der Zielgruppen von Fernbusnutzern und Nichtnutzern getrennt voneinander analysiert. Im günstigsten Fall lassen sich hier in beiden Gruppen unterschiedliche Aussagen treffen, die es ermöglichen ein genaues Bild eines Fernbusnutzers zu erhalten.

4.1.1 Zielgruppen

Zu Beginn sollen die beiden Zielgruppen der Analyse beschrieben werden. Insgesamt haben 129 den Fragebogen begonnen. Davon ist dieser 119-mal beendet worden. Daher liegt die Antwortquote bei 92,25 Prozent. Von den $n = 119$ zu betrachtenden Personen sind 63 männlich und 56 weiblich, sodass die Geschlechter relativ gleich verteilt sind. Um die Personen in die Zielgruppen Fernbusnutzer (FBN) und Nicht-Nutzer (NN) einteilen zu können, wird im Fragebogen gefragt, ob der Proband schon einmal einen Fernbus genutzt hat. Wird diese Frage mit ja beantwortet, kann die Person zu der Gruppe FBN zugeordnet werden. Wenn die Antwort nein lautet gilt der umgekehrte Schluss. Mit Hilfe der Zielgruppenanalyse soll nun ein „typischer“ Fernbusnutzer definiert werden. Die Zielgruppe der Fernbusnutzer umfasst eine Größe von $n_{FBN} = 93$. Die Gruppe der Nicht-Nutzer ist mit $n_{NN} = 26$ deutlich kleiner. In Kapitel 4.1.2 werden nun für alle erhobenen personenbezogenen Variablen die Unterschiede und Ähnlichkeiten der beiden Gruppen untersucht werden.

4.1.2 Merkmalsanalyse

Im Folgenden sollen die einzelnen erhobenen soziodemographischen Merkmale näher untersucht werden, um Unterschiede zwischen den Fernbusnutzern und Nicht-Nutzern herauszustellen. Die betrachteten Merkmale sind:

- Geschlecht
- Alter
- Beruf
- Einkommen
- Höchster Schulabschluss
- Höchste Berufsausbildung
- PKW-Verfügbarkeit
- Fernbusnutzung

Die Variablen Alter und Einkommen werden mit ordinalskalierten Klassen dargestellt. Die Merkmale Geschlecht und PKW-Verfügbarkeit sind binär. Beruf, Schulabschluss, Berufsausbildung und Fernbusnutzung werden mit Hilfe einer nominalen Skala beschrieben.

Geschlecht

Die Variable *Geschlecht* stellt mit ihren beiden Ausprägungen „männlich“ und „weiblich“ ein nominal-skaliertes Merkmal dar, welches binär beschrieben wird. Es ist festzustellen, dass die Anteile bei den Fernbusnutzern sehr ausgeglichen sind. Dabei sind 50,54 Prozent männlich und 49,46 Prozent weiblich. Die Nicht-Nutzer Gruppe besteht zu einem großen Teil aus Männern.

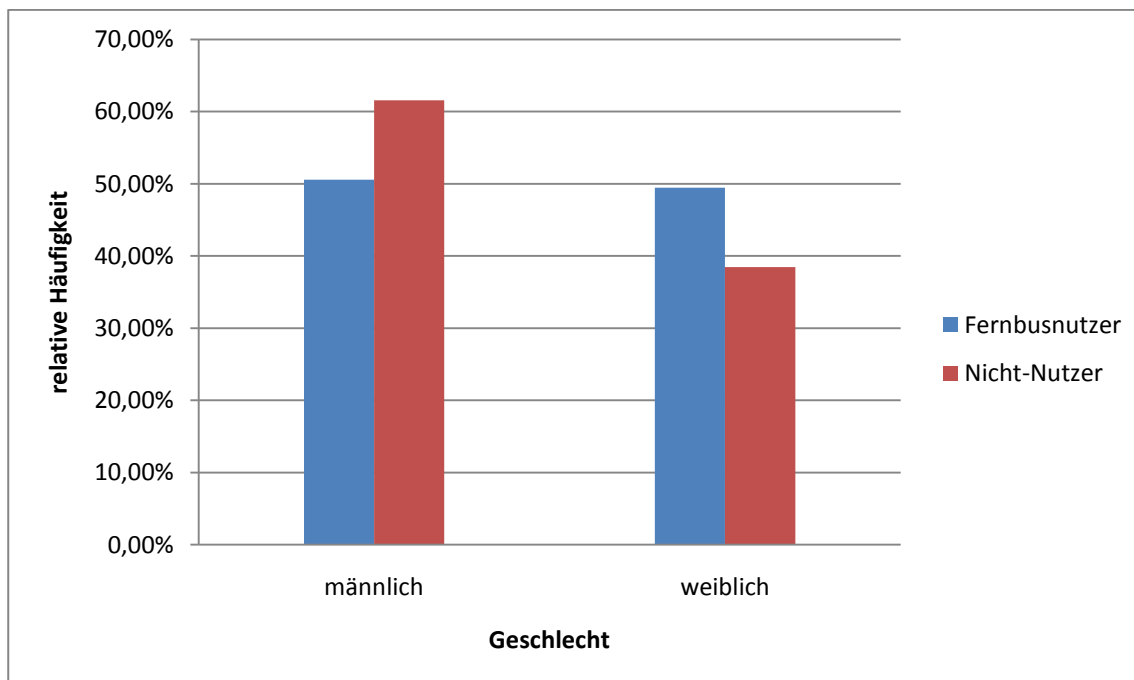


Abbildung 2: Relative Häufigkeit des Merkmals Geschlecht (e.D)

In Abbildung 2 ist die relative Häufigkeit in den beiden Gruppen für die beiden Ausprägungen „männlich“ und „weiblich“ dargestellt. Es ist festzuhalten, dass in die Nicht-Nutzer-Gruppe lediglich 26 Personen eingeteilt werden und somit der Unterschied in absoluten Zahlen geringer ausfällt. Es kann zusammenfassend also kein typisches Geschlecht für die Nutzung eines Fernbusses festgelegt werden.

Alter

Das *Alter* ist in sieben Klassen erhoben worden. Dabei handelt es sich um ein ordinal-skaliertes Merkmal und kann mit Hilfe verschiedener Lagemaße genauer untersucht werden. In Tabelle 4 ist die relative Häufigkeitsverteilung der beiden Zielgruppen für jede Altersklasse dargestellt. Über beide Gruppen gesehen, ist in jeder Altersklasse mindestens eine Person

vorhanden. Es zeigt sich in beiden Zielgruppen ein sehr hoher Anteil in der Klasse „18-25 Jahre“. Dies kann daran liegen, dass die Umfrage online durchgeführt wurde und über verschiedene soziale Netzwerke geteilt wurde. Dort sind vor allem die jüngeren Altersklassen aktiv und somit stärker vertreten. Dennoch kann bei den Nicht-Nutzern in Altersklasse „46-55 Jahre“ ein Anteil von 30,77 Prozent festgestellt werden. Des Weiteren kann festgehalten werden, dass über 90 Prozent 30 Jahre oder jünger sind. Lediglich 4,3 Prozent sind älter. Es zeigt sich also, dass vermehrt junge Personen einen Fernbus nutzen. Dennoch ist eine Nutzung von Personen, die älter als 46 Jahre sind, nicht auszuschließen.

Altersklasse	Fernbusnutzer	Nicht-Nutzer
18 – 25 Jahre	65,59 %	42,31 %
26 – 30 Jahre	30,11 %	7,69 %
31 – 35 Jahre	0,00 %	3,85 %
36 – 45 Jahre	0,00 %	3,85 %
46 – 55 Jahre	2,15 %	30,77 %
56 – 64 Jahre	2,15 %	3,85 %
65 Jahre und älter	0,00 %	7,69 %

Tabelle 4: Relative Häufigkeiten des Merkmals Alter

Da es sich bei diesem Merkmal um eine ordinal-skalierte Variable handelt, die mittels Klassen dargestellt wird, kann sie durch verschiedene Lagemaße näher beschrieben werden. Für das *Alter* sind keine genauen Werte angegeben, sodass nur die Ober- und Untergrenze einer jeden Klassen vorhanden sind. Für weitere Berechnungen muss nun die Klassenmitte jeder Klasse bestimmt werden.

$$x_C^* = \frac{x_o - x_u}{2} \quad (\text{Mosler (2009), S. 36}) \quad (4.1)$$

Da die letzte Klasse „65 Jahre und älter“ keine Obergrenze hat, wird diese auf 80 Jahre gesetzt. Dies bedeutet nicht, dass ältere Personen aus der Berechnung ausgeschlossen werden sollen. Allerdings ist die Wahrscheinlichkeit bei der geringen Anzahl an Probanden gering, dass Personen teilgenommen haben, die älter als 80 Jahre sind. Des Weiteren kann mit Hilfe der oberen und unteren Grenze die Klassenbreite ermittelt werden.

$$d_C = x_C^o - x_C^u \quad (4.2)$$

Damit können nun verschieden Lagemaße für klassierte Daten, wie z.B. Mittelwert, Median oder Modus berechnet werden. Das durchschnittliche Alter der Zielgruppen wird mittels

$$\bar{x} = \sum_{c=1}^C f_c * x_c^* \quad (\text{Mosler (2009), S. 36}) \quad (4.3)$$

berechnet. Hierbei zeigt sich für die Fernbusnutzer ein Durchschnittsalter von 25 Jahren. Der Durchschnitt der Nicht-Nutzer ist wesentlich älter. Dieser liegt bei 38 Jahren. Es lässt sich also eine Tendenz beider Zielgruppen dahingehend erkennen, dass die Nutzer von Fernbussen relativ jung sind. Eine weitere wichtige Größe ist der Modus, welcher den häufigsten Wert einer Stichprobe anzeigt. (Mosler (2009), S. 46) Dieser zeichnet sich bei klassierten Daten so aus, dass er von der Klassenmitte der Klasse mit größter Häufigkeitsdichte repräsentiert wird. Für beide Zielgruppen ist dies die Klasse „18 – 25 Jahre“. Somit liegt der Modus für beide Nutzergruppen bei 21,5 Jahren. Hier zeigt sich, dass der Großteil beider Gruppen sehr jung ist und die Ergebnisse für den Modus vom Mittelwert abweichen. Dies kann daran liegen, dass der Mittelwert anfällig für Ausreißer ist. Zuletzt kann der Median, auch Zentralwert für Klassen, berechnet werden.

$$x_{Med} = x_{C'}^u + \frac{0,5 - F_{C-1}}{f_{C'}} * d_{C'} \quad (\text{Mosler (2009), S. 58}) \quad (4.4)$$

So zeigt sich für die Nutzer ein Median von 30,34 Jahre, wohingegen der Median der Nicht-Nutzer bei 34 Jahren liegt. Hier kann die Tendenz des Mittelwertes bestätigt werden. Auch, wenn der Unterschied nicht so stark ist. Zusätzlich dazu kann die Varianz für beide Gruppen mittels

$$s^2 = \sum_{c=1}^C f_c (x_c^* - \bar{x})^2 \quad (4.5)$$

ermittelt werden (Mosler (2009), S. 59). Hier zeigt sich für die Nicht-Nutzer mit 282,08 eine deutlich höhere Varianz als bei den Fernbusnutzern mit 51,06. Mit Hilfe der Varianz kann ebenfalls die Standardabweichung berechnet werden.

$$s = \sqrt{s^2} \quad (4.6)$$

Varianz und Standardabweichung sind immer positiv und genau dann gleich Null, wenn alle Werte gleich sind. (Mosler (2009), S. 42/43) Für beide Gruppen liegt eine relativ geringe Varianz vor und somit ist die Streuung innerhalb der Gruppen klein. Der Variationskoeffizient

ist ein relatives Streuungsmaß und maßstabsunabhängig. Wenn dieser Wert größer als 0,5 ist, dann ist der Mittelwert kein guter Repräsentant der Gruppe.

$$V = \frac{s}{\bar{x}} \quad (\text{Mosler (2009), S.96}) \quad (4.7)$$

Beide Werte liegen unterhalb der 0,5 und somit sind die Mittelwerte der Zielgruppen repräsentativ für die Daten. Alle erklärten Werte und die Ergebnisse für die 25 Prozent- und 75 Prozent-Quartile sind in Tabelle 5 dargestellt.

Lagemaße	Fernbusnutzer	Nicht-Nutzer
Mittelwert	25	38
Varianz	51,06	282,08
Standardabweichung	7,15	16,80
Variationskoeffizient	0,2869	0,4479
Modus	21,5	21,5
25%-Quartil	27,69	29,14
Median	30,34	34
75%-Quartil	31,25	60,06

Tabelle 5: Lagemaße für das Merkmal Alter

Es zeigt sich also, dass der „typische“ Fernbusnutzer zwischen 18 und 30 Jahre alt ist und damit das Angebot von Fernbussen eher von der jungen Generation genutzt wird. Die Nicht-Nutzer der Stichprobe sind deutlich älter, obwohl nicht zu vergessen ist, dass auch hier der Großteil zu der jüngsten Klasse gehörte. Da aber sehr viel mehr ältere Personen zu den Nicht-Nutzern zählen, verändern sich die Ergebnisse dementsprechend.

Beruf

Das Merkmal *Beruf* ist mit sechs verschiedenen Ausprägungen nominal-skaliert. Dabei wird zwischen den allgemeinen Berufsgruppen „Schüler“, „Student“, „Azubi“, „Berufstätig“, „Rentner“ und „Arbeitssuchend“ unterschieden. Abbildung 3 zeigt die relativen Häufigkeiten in allen Berufsklassen für beide Zielgruppen. Dabei ist zu beachten, dass keine Person mit dem Status „Schüler“ an der Umfrage teilgenommen hat. Für die Fernbusnutzer zeigt sich ein

sehr hoher Anteil an Studenten. Hier kann ebenfalls davon ausgegangen werden, dass dies durch die Art der Erhebung zustande kommt. Dennoch zeigt sich vor allem bei Studenten ein größeres Interesse am Fernbusangebot, da diese mit günstigen Ticketpreisen locken. Den zweitgrößten Anteil bilden die Berufstätigen. Bei den Nicht-Nutzern liegt der Anteil von berufstätigen Personen bei 46,15 Prozent und bildet somit die größte relative Häufigkeit. Es kann vermutet werden, dass viele Berufstätige einen eigenen PKW besitzen und somit auf das Angebot des Fernbusses nicht angewiesen sind. In der Klasse der Auszubildenden ist der Anteil der Nutzer (6,45 Prozent) und Nicht-Nutzer (7,69 Prozent) fast ausgeglichen. Des Weiteren kann in den beiden verbleibenden Gruppen der Rentner und Arbeitsuchenden ein größerer Anteil an Nicht-Nutzern festgestellt werden.

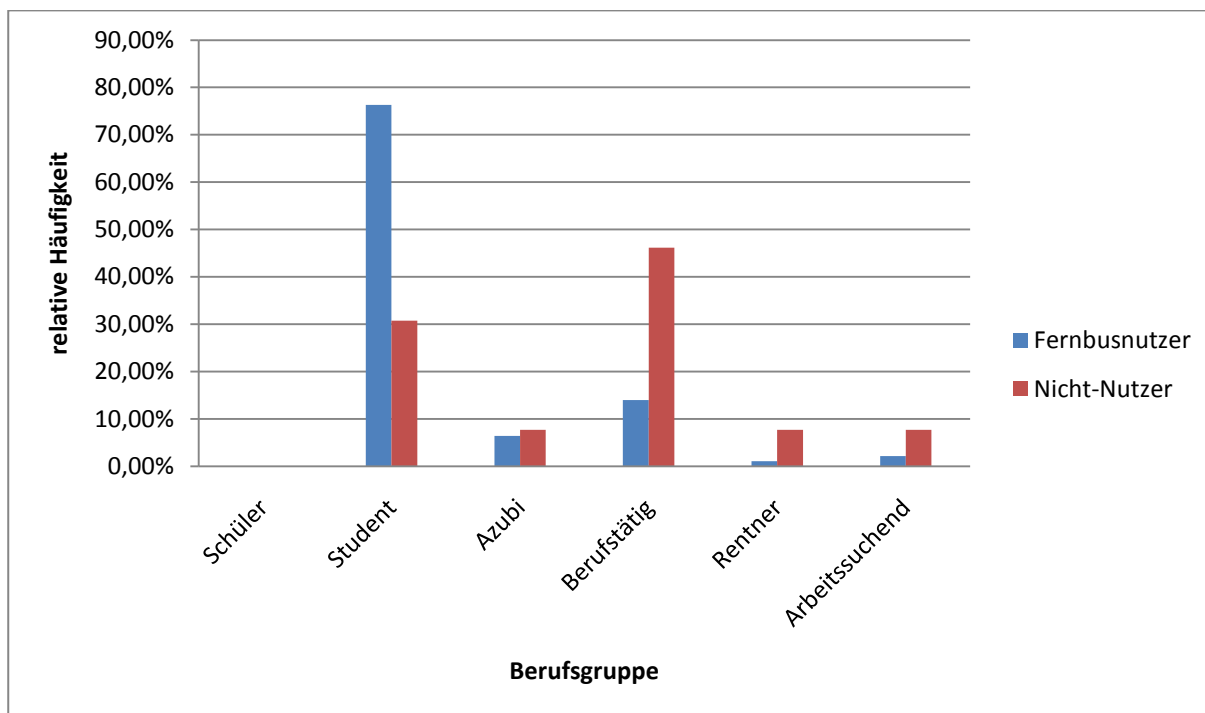


Abbildung 3: Relative Häufigkeit des Merkmals Beruf (e.D.)

So zeigt sich zusammenfassend für den Beruf, dass der „typische“ Nutzer von Fernbusangeboten Student ist.

Einkommen

Das *Einkommen* ist eine ordinal-skalierte Variablen, welche in fünf verschiedenen Einkommensklassen eingeteilt wird. Hierbei wird das monatliche Nettoeinkommen jedes Probanden erfragt. Zu Anfang können die absoluten und relativen Häufigkeiten für jede Klasse bestimmt

werden. Die relative Verteilung ist in Tabelle 6 dargestellt. Es zeigt sich, dass sowohl bei den Fernbusnutzern als auch bei den Nicht-Nutzern der größte Anteil in der Klasse „Unter 1.000€“ zu finden ist. In der Klasse „1.000€ - 2.000€“ befinden sich jeweils ca. 25 Prozent der jeweiligen Zielgruppen. Auffällig ist, dass es für die Nutzer der Fernbusse in den beiden größten Klassen keine Personen gibt. Dies ist bei den Nicht-Nutzern nicht so. Hier sind in „Über 4.500€“ noch 3,85 Prozent der befragten Nicht-Nutzer eingeordnet. Der hohe Anteil in der kleinsten Klasse lässt sich durch die große Anzahl an Studenten erklären. Diese besitzen oft ein geringeres Einkommen als alle anderen Gruppen. So lässt sich schließen, dass die Fernbusnutzer eher Geringverdiener sind.

Einkommensklassen	Fernbusnutzer	Nicht-Nutzer
Unter 1.000€	72,04 %	50,00 %
1.000€ – 2.000 €	23,66 %	26,92 %
2.000€ – 3.000€	4,30 %	11,54 %
3.000€ – 4.500€	0,00 %	7,69 %
Über 4.500€	0,00 %	3,85 %

Tabelle 6: Relative Häufigkeiten des Merkmals Einkommen (e.D.)

Das Merkmal *Einkommen* kann des Weiteren auch wie das *Alter* mit verschiedenen Lagemaßen dargestellt werden. Da es sich hier um Einkommensklassen handelt und somit keine genauen Werte sondern nur obere und untere Grenze einer Klasse vorhanden sind, wird wie bei dem Merkmal *Alter* vorgegangen. Die Ergebnisse der Berechnungen sind in Tabelle 7 zu finden.

Lagemaße	Fernbusnutzer	Nicht-Nutzer
Mittelwert	822,08	1.451,44
Varianz	304.534,88	1.697.373,16
Standardabweichung	551,85	1.302,83
Variationskoeffizient	0,6713	0,8976
Modus	499,50	499,50
25 %-Quartil	1.345,68	1.498,5
Median	1.692,36	1.998,00
50 %-Quartil	2.123,98	2.926,75

Tabelle 7: Lagemaße für das Merkmal Einkommen

Im Durchschnitt hat ein Fernbusnutzer 822,08 Euro im Monat zur freien Verfügung. Die Nicht-Nutzer besitzen mit 1.451,44 Euro im Monat ein deutlich höheres Nettoeinkommen. Dies bestätigt auch der Median. 50 Prozent der Nutzer besitzen nicht mehr als 1.692,36 Euro im Monat, während 50 Prozent der Nicht-Nutzer nicht mehr als 1.998 Euro im Monat zur Verfügung haben. Es zeigt sich also, dass Nutzer von Fernbusangeboten weniger Geld im Monat haben als die andere Zielgruppe. Wie auch bei dem *Alter* ist der Modus der Gruppen gleich. Dies liegt wieder daran, dass der größte Anteil der Probanden in die Gruppe „Unter 1.000€“ eingeordnet werden. Des Weiteren ist festzustellen, dass die Varianz und somit auch die Standardabweichung für beide Zielgruppen sehr hoch sind. Dies hat ebenfalls Auswirkungen auf den Variationskoeffizienten. Dieser liegt für beide Gruppen bei deutlich über 0,5. Somit kann festgestellt werden, dass die Mittelwerte die Daten nicht gut repräsentieren. So zeigt sich zusammenfassend, dass ein Fernbusnutzer eher weniger verdient. Durch die niedrigen Ticketpreise ist dieses Verkehrsmittel sehr attraktiv für diese Zielgruppe.

Höchster Schulabschluss

Das Merkmal *Höchster Schulabschluss* besitzt vier Ausprägungen. Diese werden in „Hauptschule“, „Realschule“, „Abitur“ und „ohne Abschluss“ unterschieden. Abbildung 4 zeigt die Ergebnisse der relativen Häufigkeit für jede Ausprägung in den beiden Zielgruppen.

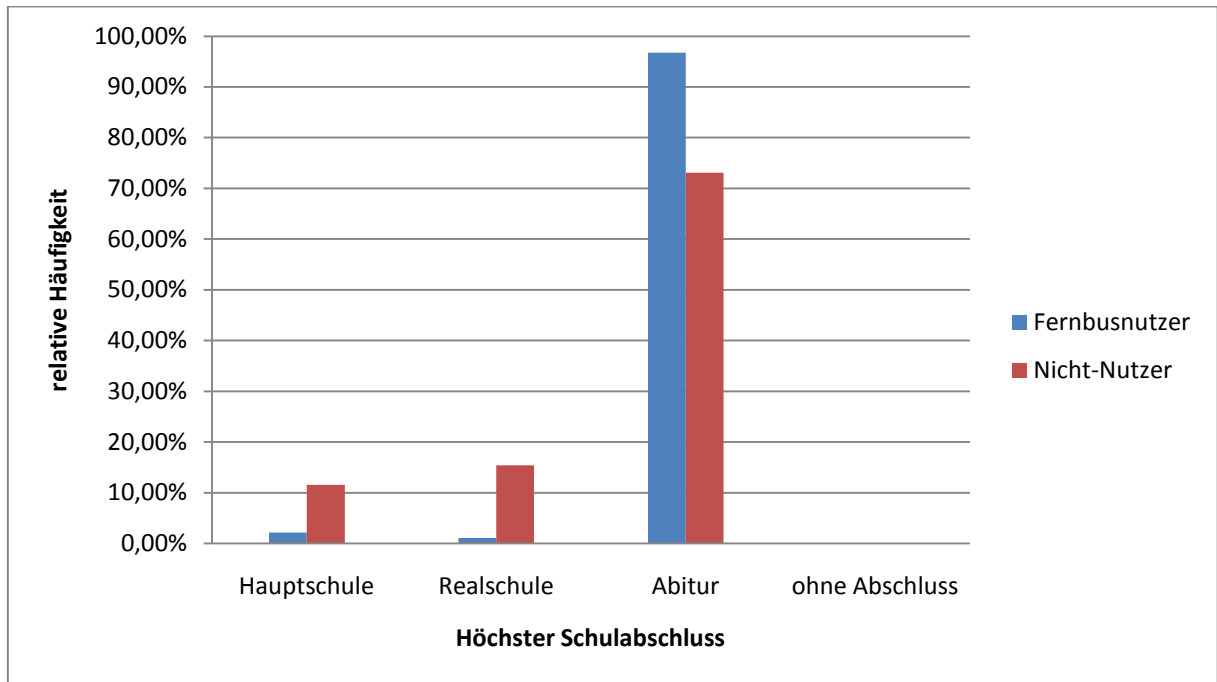


Abbildung 4: Relative Häufigkeit des Merkmals *Höchster Schulabschluss* (e.D.)

Es zeigt sich, dass in beiden Zielgruppen der größte Anteil der Probanden das Abitur besitzt und dementsprechend gut gebildet ist. Bei den Fernbusnutzern liegt der Anteil bei 96,77 Prozent. Nur ein sehr geringer Anteil von 3,23 Prozent wird auf die restlichen Gruppen verteilt. Dabei ist zu beachten, dass alle Probanden einen Schulabschluss besitzen. Dies kann auch aus dem Merkmal Beruf geschlossen werden, da keiner der Teilnehmer den Berufsstatus „Schüler“ innehat. Bei den Nicht-Nutzern besitzen 73,08 Prozent das Abitur. Der zweitgrößte Anteil (15,38 Prozent) hat einen Realschulabschluss und die kleinste Gruppe bilden Personen mit einem Hauptschulabschluss mit 11,54 Prozent. So zeigt sich, dass Nutzer von Fernbusangeboten ein sehr gutes Bildungsniveau besitzen.

Höchste Berufsausbildung

Neben der höchsten Schulausbildung ist auch die *Höchste Berufsausbildung* im Fragebogen erfragt worden. Diese Variable ist ebenfalls nominal-skaliert und wird in die vier Klassen

„Lehre/Berufsfachschule“, „Meister-/Fachschule“, „Hoch-/Fachhochschulabschluss“ und „(noch) ohne Abschluss“ eingeteilt. Bei den Nutzern des Fernbusangebots zeigt sich ein Anteil von 50,54 Prozent, die noch keinen Abschluss besitzen. Dies kann an der sehr hohen Anzahl der Studenten liegen, die noch auf den Abschluss ihrer Berufsausbildung hinarbeiten. Des Weiteren besitzen 37,63 Prozent einen Hoch-/Fachhochschulabschluss. Somit zeigt sich unter den Fernbusnutzern ein hohes Bildungsniveau. Die Anteile für die beiden anderen Klassen belaufen sich auf 11,83 Prozent.

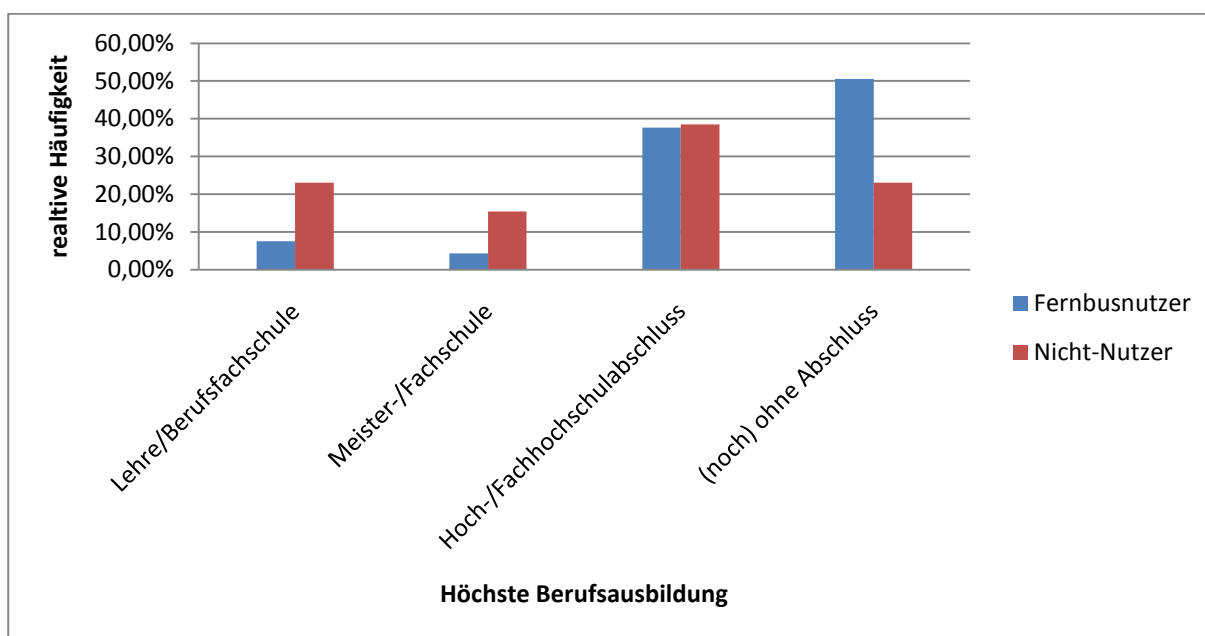


Abbildung 5: Relative Häufigkeit des Merkmals Höchste Berufsausbildung (e.D.)

In Abbildung 5 sind ebenfalls die relativen Häufigkeiten für die Nichtnutzer in den einzelnen Ausprägungen dargestellt. Anders als in der Zielgruppe der Nutzer haben die meisten Probanden einen Hoch-/Fachhochschulabschluss. Die Ausprägungen „Lehre/Berufsfachschule“ und „(noch) ohne Abschluss“ bilden mit jeweils 23,08 Prozent die zweitgrößten Anteile. Genau wie bei den Fernbusnutzern ist der Anteil der Personen, die eine Meister-/Fachschule besucht haben, am geringsten. Zusammenfassend zeigt sich, dass der Fernbusnutzer eine gute Ausbildung absolviert hat oder noch absolviert und sich somit kleine Unterschiede im Bildungsniveau der Zielgruppen feststellen lassen.

PKW-Verfügbarkeit

Ein weiterer Einfluss, der auf die Zielgruppenzugehörigkeit schließen lässt, ist die *PKW-Verfügbarkeit*. Es kann angenommen werden, dass Personen mit eigenem PKW weniger mit Fernbussen reisen, als Personen, die keinen PKW besitzen, da sie meist flexibler und schneller unterwegs sind. Ist kein PKW vorhanden, muss für eine Reise in andere Städte ein öffentliches Verkehrsmittel, wie der Fernbus oder die Bahn, genutzt werden.

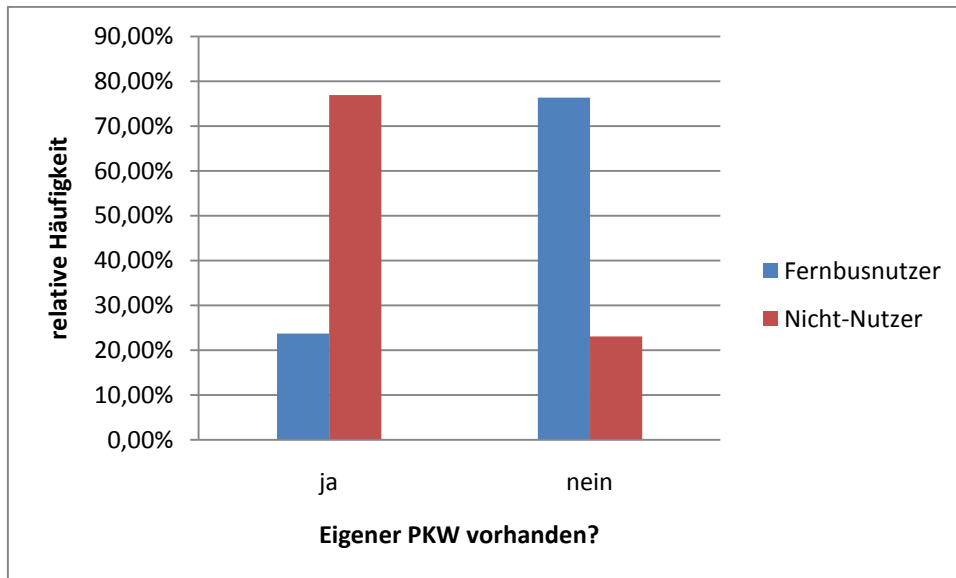


Abbildung 6: Relative Häufigkeit der PKW Verfügbarkeit (e.D.)

Abbildung 6 zeigt die Verteilung in den beiden Zielgruppen für die *PKW-Verfügbarkeit*. Auffallend ist die umgekehrte Verteilung für die Gruppen. Während 76,34 Prozent der Fernbusnutzer keinen eigenen PKW besitzen, haben 76,92 Prozent der Nicht-Nutzer einen PKW vor der Haustür stehen. Allerdings lässt sich feststellen, dass auch Personen mit PKW das Angebot des Fernbusses nutzen. Dies kann an den günstigen Ticketpreisen liegen. 23,08 Prozent der Nicht-Nutzer besitzen ebenfalls keinen PKW. Hier kann vermutet werden, dass bevorzugt mit anderen Verkehrsmitteln gereist wird.

Fernbusnutzung

Ein weiterer interessanter Aspekt, der nur für die Zielgruppe der Nutzer erhoben werden kann, ist die Nutzung des Fernbusses. Dabei geht es um die Häufigkeit von Fernbusfahrten der Personen. Ein Proband, der in der Gruppe der Nicht-Nutzer ist, kann hier nur mit der Antwort „nie“ antworten und somit wird dies nicht berücksichtigt. Tabelle 8 zeigt die relative Häufig-

keitsverteilung auf sieben verschiedene Ausprägungen, wobei die Ausprägung „nie“ hier nicht gezeigt wird, da sie bei einem Nutzer nicht gewählt werden kann.

Nutzungshäufigkeit	Fernbusnutzer
1 - 2x im Jahr	31,18 %
1 - 2x in 6 Monaten	33,33 %
1 - 2x in 3 Monate	11,83 %
1 - 2x im Monat	7,53 %
1 - 2x in der Woche	1,08 %
seltener	15,05 %
öfter	0,00 %

Tabelle 8: Relative Häufigkeiten des Merkmals Fernbusnutzung (e.D.)

79,56 Prozent der Fernbusnutzer fahren ein bis zweimal in sechs Monaten mit dem Fernbus. Dies könnten Besuche bei Freunden oder Familie sein, die an Wochenende getätigt werden. Allerdings zeigt sich auch, dass 1,08 Prozent jede Woche den Fernbus nutzen. Eine Möglichkeit hierfür wäre das Pendeln von Paaren am Wochenende.

Es zeigt sich also, dass der „typische“ Fernbusnutzer zwischen 18 und 30 Jahre alt ist und noch studiert. Das Einkommen ist dadurch eher gering. Dennoch ist der Fernbusnutzer sehr gebildet und besitzt keinen eigenen PKW.

4.2 Korrelationsanalyse

Anschließend wird die Korrelation zwischen den soziodemographischen Variablen untersucht. Zuerst werden die nominal-skalierten Variablen Beruf, Schulabschluss und Berufsausbildung mittels Rangkorrelationskoeffizienten nach Spearman analysiert. Zusätzlich werden die Variablen Beruf und Alter mit der gleichen Formel untersucht.

$$r_{XYI}^R = 1 - \frac{6 \sum_{i=1}^n (R_X(x_i) - R_{YI}(y_i))^2}{n(n^2 - 1)} \quad (4.8)$$

(Mosler (2009), S. 178)

Dabei kann zwischen den drei untersuchten Variablen eine nahezu perfekte Korrelation nachgewiesen werden. Dies gilt auch für die Variablen Alter und Beruf. Anschließend kann mittels Chi-Quadrat-Unabhängigkeitstest geprüft werden, ob es einen Zusammenhang zwischen dem Alter der Person und ihrer Zugehörigkeit zu den Fernbusnutzern oder Nichtnutzern besteht. Dabei zeigt sich, dass die Variable Alter und die Zielgruppenzugehörigkeit voneinander abhängig sind. Dies kann in nachfolgender Tabelle 9 nachvollzogen werden.

h_{ip}	18-25	26-30	31-35	36-45	46-55	56-64	65+	h_p
$y=1$	61	28	0	0	2	2	0	93
$y=0$	11	2	1	1	8	1	2	26
h_i	72	30	1	1	10	3	2	119

Realisierung

h_{ip} erwartet	17-25	25-35	35-45	36-45	46-55	56-64	65+	h_p
$y=1$	56,3	23,4	0,8	0,8	7,8	2,3	1,6	93
$y=0$	15,7	6,6	0,2	0,2	2,2	0,7	0,4	26
h_i	72	30	1	1	10	3	2	119

Testvariable	40,21
Freiheitsgrade m	7
χ^2 -Wert	14,07

> Variablen sind abhängig!

Tabelle 9: χ^2 -Unabhängigkeitstest von Alter und Zielgruppenzugehörigkeit

Die Korrelationsmatrix mit allen untersuchten personenbezogenen Merkmalen findet sich im Anhang (siehe A.3). Auch wenn keine metrischen Variablen vorhanden sind, zeigen sich einige Zusammenhänge wie z.B. zwischen Alter und Beruf (0,63) oder Alter und Einkommen (0,50). Ebenfalls findet sich eine positive Korrelation zwischen Alter und der PKW-Verfügbarkeit von 0,43. So zeigt sich, dass ältere Personen eher ein eigenes Kraftfahrzeug besitzen als junge Menschen. Des Weiteren besteht zwischen der PKW-Verfügbarkeit und der Zielgruppenzugehörigkeit ein negativer Zusammenhang. Dies lässt sich dadurch erklären, dass der Fernbus eher genutzt wird, wenn kein PKW vorhanden ist.

4.3 Auswahlbasierte Conjoint-Analyse

Im zweiten Teil der Analyse wird eine auswahlbasierte Conjoint-Analyse durchgeführt. Dazu werden die Auswahlentscheidungen der Probanden für verschiedene Fernbusstrecken analy-

siert. Wie im Fragebogen (Anhang A.1) ersichtlich, wurden jeweils vier lange Strecken und vier kurze Strecken betrachtet. Die Merkmale beschränken sich auf Reisezeit, Kosten und Vorhandensein eines extra ÖV-Tickets. Das Erhebungsdesign für die langen und kurzen Strecken ist jeweils gleich. Mit Hilfe eines Fragebogens, welcher im Internet freigeschaltet wird, werden die Testpersonen und Ergebnisse ermittelt. Für das Merkmal Reisezeit gibt es die Ausprägungen „kurze Reisezeit“ und „lange Reisezeit“ und für die Kosten sind die beiden Ausprägungen „niedrige Kosten“ und „hohe Kosten“ vorhanden, welche jeweils binär dargestellt werden. Da ein extra ÖV-Ticket nur in Variante 2 vorhanden ist, wird dies aus der Analyse ausgeschlossen. Variante 1 im Fragebogen stellt immer die Situation mit Haltestellen in der Innenstadt dar, wohingegen Variante 2 die Haltestelle in Routennähe simuliert. Da insgesamt acht mögliche Entscheidungssituationen vorgestellt werden und die Probandenanzahl bei 119 liegt, gibt es insgesamt $n_a = 952$ Entscheidungssituationen, wobei 476 auf die langen Strecken und 476 auf die kurzen Strecken fallen. Im Folgenden wird das Analyseverfahren für die Conjoint-Analysen, wie in Kapitel 3.2.2 beschrieben, angewendet.

4.3.1 Lange Strecken

Zuerst werden die ersten vier Schritte (Gestaltung der Stimuli, Gestaltung der Auswahl-situation, Spezifikation des Nutzenmodells und Spezifikation eines Auswahlmodells) des Auswahlverfahrens gekürzt dargestellt. Es liegen die beiden Merkmale Kosten und Reisezeit mit jeweils zwei Ausprägungen vor. Allerdings werden diese nicht beliebig kombiniert, sondern systematisch für die beiden betrachteten Varianten zusammengesetzt (siehe Tabelle 10).

Variante	Kosten	Reisezeit
Haltestelle in der Stadt (1)	Niedrige Kosten	Lange Reisezeit
Haltestelle in Routennähe (2)	Hohe Kosten	Kurze Reisezeit

Tabelle 10: Darstellung der Varianten bei langen Strecken

Für die langen Strecken, bei denen mehrere Städte angefahren werden, bedeutet dies, dass Variante 1 mit Haltestelle in der Stadt günstiger ist als Variante 2, aber die Reisezeit länger ist. Bei allen vier ausgewählten Strecken ist dies der Fall. Es ist keine *None-Option* vorhanden, sodass der Proband pro Auswahl-situation zwischen zwei Alternativen entscheiden kann. Für

die Analyse wird ein Gesamtnutzen-Modell verwendet, welches wie folgt beschrieben werden kann:

$$u_{kr} = \sum_{j=1}^J \sum_{m=1}^{M_j} b_{jm} * x_{jmk_r} \quad (4.9)$$

So wird der Nutzen der Alternative k in der Auswahl-situation r berechnet (Backhaus (2015), S. 189). Des Weiteren wird für das Auswahlmodell ein binäres Logit-Choice-Modell verwendet, wobei die Auswahlwahrscheinlichkeiten nur durch die Differenzen der einzelnen Nutzenwerte bestimmt werden. Nach der Spezifikation der Modelle werden die Nutzenwerte geschätzt und die CBCA durchgeführt. Um die Startwerte der Teilnutzenwerte zu ermitteln, werden die einzelnen Häufigkeiten der Ausprägungen berechnet und durch die Anzahl der Auswahl-situationen geteilt. Für jedes Merkmal wird der Teilnutzenwert einer Ausprägung auf null gesetzt und der Teilnutzen der anderen Ausprägung besteht aus der Differenz der beiden Teilnutzenwerte einer Variable. (siehe Tabelle 11)

	Hohe Kosten	Niedrige Kosten	Lange Reisezeit	Kurze Reisezeit
Absolute Häufigkeit	443	33	33	443
Auswahl-situationen	476	476	476	476
Anteil	0,9307	0,0693	0,0693	0,9307
Startwert	0,8613	0,0000	- 0,8613	0,0000

Tabelle 11: Ermittlung Startwerte der langen Strecken

Mit Hilfe der Startwerte kann für jede Alternative in einer Situation der Gesamtnutzen mittels Formel (4.9) berechnet werden. Anschließend werden mit Formel (3.12) die Auswahlwahrscheinlichkeiten ermittelt und der LogLikelihood-Wert über alle Situationen und Alternativen für jeden Probanden berechnet. Dieser wird mit der Veränderung der Startwerte maximiert (Formel (3.10)). Nach der Schätzung können folgende Ergebnisse festgestellt werden:

Ausprägungen	Hohe Kosten (b ₁₁)	Niedrige Kosten (b ₁₂)	Lange Reisezeit (b ₂₁)	Kurze Reisezeit (b ₂₂)
Schätzwerte	1,2985	0,0000	- 1,2986	0,0000

Tabelle 12: Schätzwerte der Teilnutzen nach ML-Schätzung der langen Strecken

In Tabelle 12 sind die geschätzten Teilnutzen nach der Maximum-Likelihood-Schätzung dargestellt. Es zeigt sich, dass die Ausprägung „hohe Kosten“ einen größeren Nutzen hat als „niedrige Kosten“. Des Weiteren ist der Teilnutzen der langen Reisezeit geringer als der der kurzen Reisezeit. Dies bedeutet, dass hohe Kosten und eine kurze Reisezeit, wie in der Situation mit Haltestellen in Routennähe, der Variante mit niedrigeren Preisen aber einer längeren Reisezeit vorgezogen werden. Diese sind in Abbildung 7 noch einmal grafisch dargestellt.

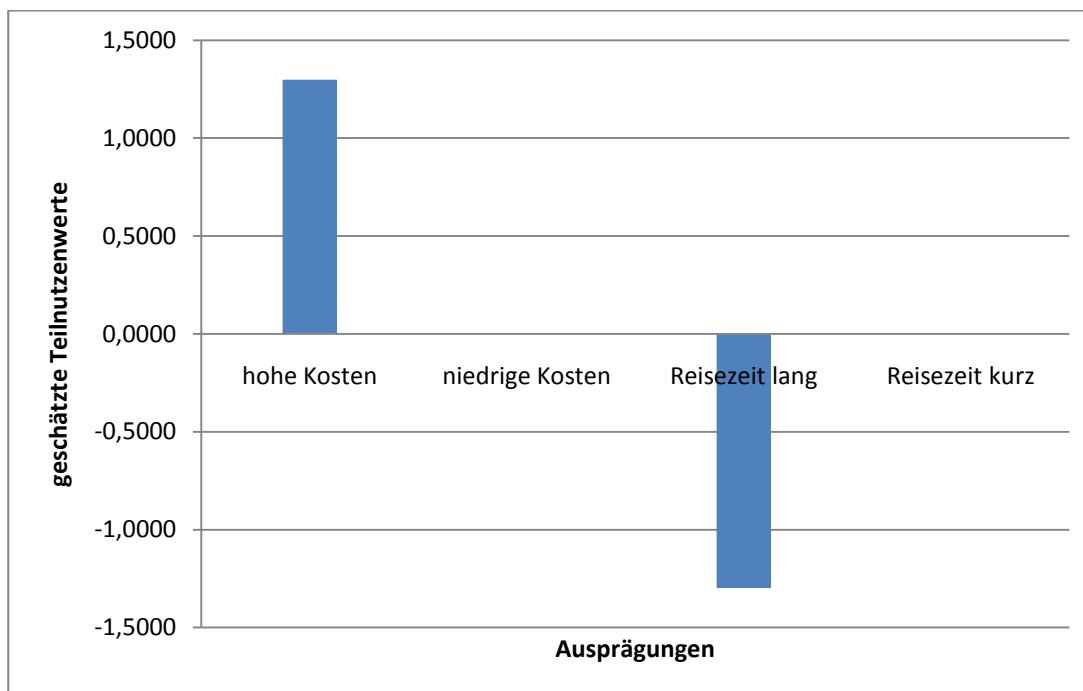


Abbildung 7: Geschätzte Teilnutzenwerte (e.D.)

Aus den Teilnutzenwerten können nun die relative Wichtigkeit der einzelnen Merkmale und die Gesamtnutzen für die Varianten errechnet werden. Für die relative Wichtigkeit einer Variable müssen die Teilnutzen zentriert werden und der Range berechnet werden. Dabei wird der Range eines Merkmals mit der Summe beider Merkmale in ein Verhältnis gesetzt, um so zu ermitteln, welche Eigenschaft für die Probanden wichtiger ist.

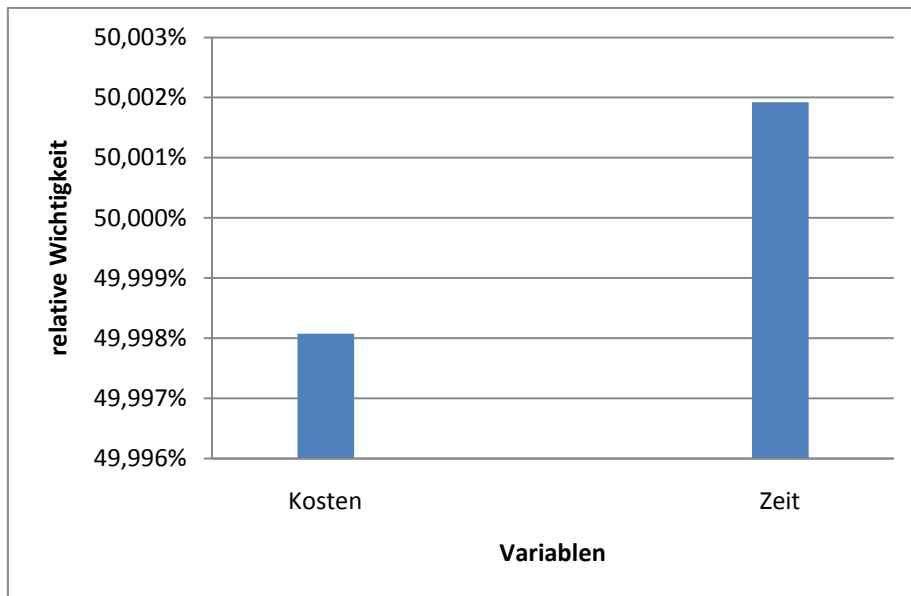


Abbildung 8: Relative Wichtigkeit der Eigenschaften der langen Strecken (e.D.)

So zeigt sich in Abbildung 8, dass die Reisezeit wichtiger für die Kunden ist, als die Kosten. Dabei ist allerdings zu beachten, dass es sich hier nur um eine Differenz von 0,004 Prozent handelt. Dieses Ergebnis bestätigt die geschätzten Teilnutzenwerte, da hier eher die Variante mit höheren Kosten aber einer geringen Reisezeit einen größeren Nutzen hat. Der Gesamtnutzen für jede Variante wird nun durch Addition der Teilwerte berechnet. Für die Variante 1 (Haltestelle in der Stadt) bedeutet dies:

$$b_{12} + b_{21} = u_1 \quad (4.10)$$

Werden die Teilnutzenwerte eingesetzt, kann folgender Gesamtnutzen für die Haltestellen in der Stadt ermittelt werden:

$$0,0000 + (-1,2985) = -1,2985$$

Der Gesamtnutzen für die Haltestellen in Routennähe liegt bei 1,2985 und ist somit größer. Das bedeutet, dass für lange Strecken mit vielen Zwischenstopps eine Haltestelle in Routennähe für die Kunden mehr Nutzen hat, da viel Zeit eingespart wird. Auch bei höheren Kosten ist diese Wahl attraktiver. Dieses Ergebnis ist in Abbildung 9 grafisch dargestellt.

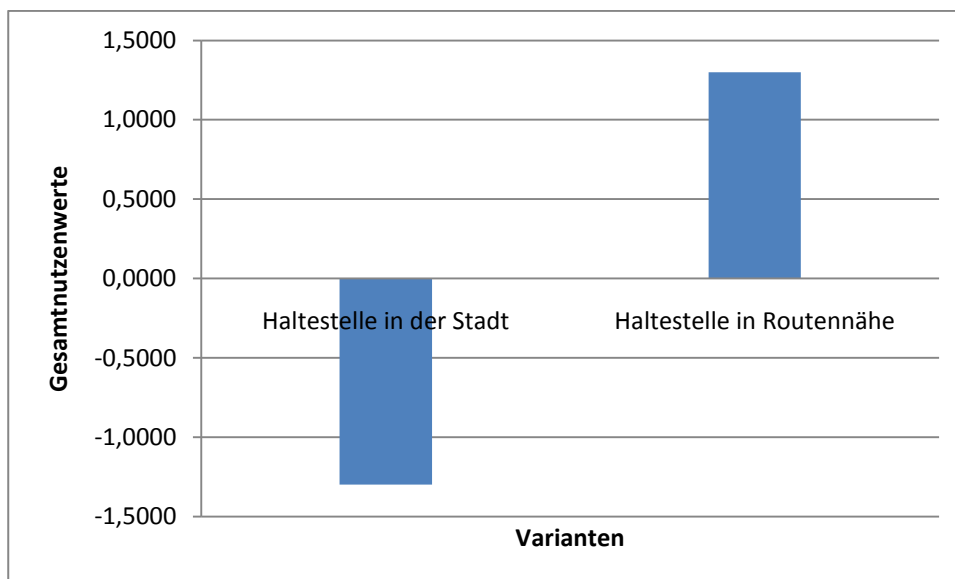


Abbildung 9: Gesamtnutzenwerte für die langen Strecken (e.D.)

Anschließend wird eine Güteprüfung durchgeführt. Dabei kann das Modell global mit Hilfe des Likelihood-Ratio-Tests oder der Trefferquote geprüft werden. Beim Likelihood-Ratio-Test wird das gesamte Modell mit dem Nullmodell verglichen. Bei einem Nullmodell sind alle Teilnutzen auf null gesetzt. In Formel (3.14) ist die Berechnung des *LLR*-Wertes dargestellt.

Güteprüfung	
LL-Wert	- 119,9027
LL0-Wert	- 329,9381
LLR-Wert	- 420,0706
p-Wert (Chi²-Verteilung)	0,0000
McFadden-R²	63,33%
Trefferquote	93,07%

Tabelle 13: Güteprüfung des Modells (Lange Strecken)

In Tabelle 13 sind alle nötigen Werte für eine globale Güteprüfung abgebildet. Dabei zeigt sich, dass das geschätzte Modell statistisch hochsignifikant ist, da der p-Wert mit 0,0000 sehr klein ist. Des Weiteren kann festgestellt werden, dass der Wert für das McFadden-R²-Gütemaß bei 63,33 Prozent liegt und somit zufriedenstellend ist. Im Allgemeinen gelten hier

Werte zwischen 20 und 40 Prozent als gute Anpassung. Allerdings kann der Wert nicht als erklärter Anteil der Streuung gesehen werden (Backhaus (2015), S. 210). Die Trefferquote des Modells liegt hier bei 93,07 Prozent. Hierbei wird angegeben in wie viel Prozent der Fälle die gewählte Alternative richtig vorhergesagt wird. Für einen „Treffer“ wird die maximale Wahrscheinlichkeit für die Alternative gewertet. Es zeigt sich also, dass das Modell für die langen Strecken einen sehr guten Modellfit besitzt, statistisch hochsignifikant ist und sehr genaue Vorhersagen für die Wahl der Varianten treffen kann. Anschließend an das globale Gütemaß, welches das gesamte Modell betrachtet, können auch die einzelnen Koeffizienten auf ihre Signifikanz untersucht werden.

Ausprägung	Hohe Kosten	Lange Reisezeit
Schätzwert (b_{jm})	1,2985	- 1,2985
LL0_j-Wert (reduziert)	- 157,7178	- 157,7250
LL-Wert (vollständig)	- 119,9072	- 119,9072
Differenz	37,8150	37,8222
LLR_j	75,6301	75,6445
p-Wert	0,0000	0,0000
Signifikanz	Stark signifikant	Stark signifikant

Tabelle 14: Prüfung der Koeffizienten der langen Strecken

In Tabelle 14 sind die Ergebnisse der Koeffizientenprüfung dargestellt. Dabei wird der Likelihood-Ratio-Test angewendet, um die Signifikanz der einzelnen Koeffizienten zu prüfen. Das vollständige Modell soll dabei mit den verschiedenen reduzierten Modellen (eins für jeden Koeffizienten) verglichen werden.

$$LLR_j = -2 * (LL0_j - LL_V) \quad (\text{Backhaus (2015), S. 211}) \quad (4.11)$$

So zeigt sich für beide Variablen, dass sie einen starken Einfluss auf das Modell haben und somit sehr signifikant für die Modellierung sind. Die Differenz der beiden LogLikelihood-Werte zeigt an, dass das Weglassen einer Variable einen starken Einfluss auf das Modell hat und somit von großer Bedeutung ist. Des Weiteren zeigt sich, dass das Merkmal der langen

Reisezeit eine marginal größere Differenz aufweist, sodass diese den größeren Einfluss auf die Modellierung und die Auswahlentscheidung der Probanden hat.

Die Untersuchung wird ebenfalls mit einer größeren Aufspaltung der Zeitvariablen durchgeführt. Dabei werden nicht nur die Kosten und die gesamte Reisezeit betrachtet, sondern die Reisezeiten zur Haltestelle und von der Haltestelle zum eigentlichen Zielort werden gesondert untersucht, sodass nun vier Variablen betrachtet werden, welche jeweils zwei Ausprägungen besitzen. Dabei zeigen sich ähnliche Ergebnisse wie zuvor. Dennoch zeigen sich bei den langen Reisezeiten zu und von den Haltestellen nun positive Nutzenwerte. Der Teilnutzenwert der langen Reisezeit im Bus ist weiterhin negativ. Die Ergebnisse sind in Tabelle 15 zu sehen.

Eigenschaft	Ausprägung	Teilnutzen	rel. Wichtigkeit
Kosten	Hoch	0,6492	25,37 %
	Niedrig	0,0000	
Zeit (bis Haltestelle)	Lang	0,6492	25,37 %
	Kurz	0,0000	
Zeit (im Bus)	Lang	- 0,6493	23,90 %
	Kurz	0,0000	
Zeit (von Haltestelle)	Lang	0,6492	25,37 %
	Kurz	0,0000	

Tabelle 15: Teilnutzen lange Strecke mit Aufspaltung der Zeit

Dabei zeigt sich, dass die Reisezeit im Bus die geringste Wichtigkeit für das Modell besitzt. Ebenfalls kann dadurch festgestellt werden, dass die Variante 2 mit den Haltestellen in Routennähe einen größeren Gesamtnutzenwert besitzen als Situation 1, welche die Haltestellen im Stadtzentrum darstellt. Diese Situation ist in Abbildung 10 dargestellt.

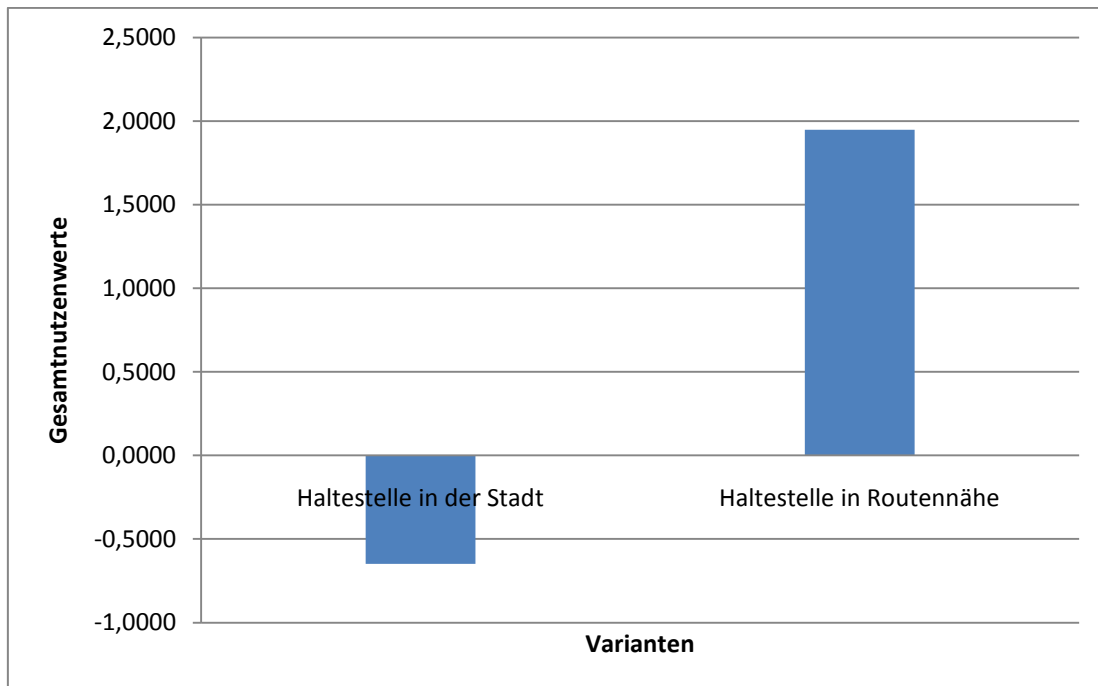


Abbildung 10: Gesamtnutzenwerte der langen Strecke mit Aufspaltung der Reisezeit (e.D.)

So können die Ergebnisse der vorangegangenen Analyse bestätigt werden. Für lange Strecken sind Haltestellen in Routennähe mit einer direkten Anbindung an den öffentlichen Personenverkehr von höherem Nutzen, weil viel Reisezeit eingespart wird.

Mit Hilfe der Ergebnisse können die Wahrscheinlichkeiten für eine Wahl der jeweiligen Situation bestimmt werden.

$$prob(1|2) = \frac{e^{u_1}}{e^{u_1} + e^{u_2}} = \frac{e^{-1,2986}}{e^{-1,2986} + e^{1,2985}} = 0,0693 \quad (4.12)$$

$$prob(2|1) = \frac{e^{1,2985}}{e^{1,2985} + e^{-1,2986}} = 0,9307 \quad (\text{Backhaus (2015), S.213})$$

So zeigt sich, dass die Variante 2 mit Haltestellen in Routennähe zu 93,07 Prozent gegenüber der Variante 1 von den Probanden bevorzugt werden würden, wenn beide Alternativen zur gleichen Zeit angeboten werden würden.

Ein Problem, dass bei der CBCA auftreten kann, ist die Überschätzung des starken Produktes und die Unterschätzung des schwächeren Produktes. Um dem vorzubeugen, kann ein Rationalitätsparameter ρ eingeführt (siehe Formel 3.7). Dieser ist üblicherweise kleiner als eins. Liegt ρ bei 0,5 können die Wahrscheinlichkeiten wie folgt berechnet werden:

$$prob(1|2) = 0,2144$$

$$prob(2|1) = 0,7856$$

Auch nach der Einbeziehung des Rationalitätsparameters, der die reale Marktsituation besser beschreiben soll, entscheiden sich 78,56 Prozent für die Variante 2 und bevorzugen damit geringe Reisezeit zu einem höheren Preis. Eine Disaggregation der Nutzenwerte ist hier nicht sinnvoll, da eine weitestgehend homogene Nutzenvorstellung der Probanden vorliegen sollte.

4.3.2 Kurze Strecken

Die Gestaltung und Spezifikation für das Analyseverfahren der kurzen Strecke ist genau wie in Kapitel 4.3.1 eingestellt. Lediglich in der Gestaltung der Varianten ist für die Reisezeit eine Änderung aufgetreten. Da nun keine zusätzlichen Städte auf der Strecke angefahren werden, sind die Reisezeiten im Bus für beide Varianten gleich. Lediglich die Zeiten für die Fahrten zu den Fernbushaltestellen und in der Zielstadt von der Haltestelle zum Zielort unterscheiden sich. Für die Situation, dass sich die Haltestelle in Routennähe befindet, sind diese beiden Zeiten länger als für die Variante 1 mit Haltestellen im Stadtzentrum. Die genauen Ausprägungen für die Varianten können in Tabelle 16 eingesehen werden.

Variante	Kosten	Reisezeit
Haltestelle in der Stadt (1)	Niedrige Kosten	Kurze Reisezeit
Haltestelle in Routennähe (2)	Hohe Kosten	Lange Reisezeit

Tabelle 16: Darstellung der Varianten bei kurzen Strecken

Des Weiteren können die Startwerte wie bei den langen Strecken berechnet werden. In Tabelle 17 sind die Startwerte und die Schätzwerte, die mittel der Maximum-Likelihood-Schätzung ermittelt werden, dargestellt.

	Hohe Kosten (b ₁₁)	Niedrige Kosten (b ₁₂)	Lange Reisezeit (b ₂₁)	Kurze Reisezeit (b ₂₂)
Absolute Häufigkeit	96	381	96	381
Auswahlsituationen	476	476	476	476
Anteil	0,2017	0,8004	0,2017	0,8004
Startwert	- 0,5987	0,0000	- 0,5987	0,0000
Schätzwerte	- 0,6829	0,0000	- 0,6892	0,0000

Tabelle 17: Ermittlung Startwerte und Schätzwerte nach ML-Schätzung der kurzen Strecken

Hierbei zeigt sich, dass anders als bei den langen Strecken, die Ausprägung „hohe Kosten“ einen negativen Schätzwert besitzt. Das bedeutet, dass der Nutzen der niedrigen Kosten höher liegt als der Nutzen der hohen Kosten. Bei der Variable Reisezeit ist für beide Ausprägung die gleiche Richtung wie in 4.3.1 zu erkennen. Die langen Reisezeiten haben für den Kunden einen geringen Nutzen als die kürzere Reisezeit. Werden diese Ergebnisse unter Berücksichtigung der Varianten betrachtet, zeigt sich, dass die Variante 1, mit niedrigen Ticketpreisen und kurzen Reisezeiten, einen höheren Nutzen besitzt als die Variante 2, die ihre Haltestellen in Routennähe haben.

Eigenschaften	Ausprägung	Teilnutzen	zentriert	Range	Rel. Wichtigkeit
Kosten	Hoch	- 0,6892	- 0,3446	0,6892	50,00 %
	Niedrig	0,0000	0,3446		
Reisezeit	Lang	- 0,6892	- 0,3446	0,6892	50,00 %
	kurz	0,0000	0,3446		

Tabelle 18: Relative Wichtigkeit der Variablen

In Tabelle 18 ist die relative Wichtigkeit der einzelnen Merkmale dargestellt. Es zeigt sich, dass Kosten und Reisezeit mit jeweils 50 Prozent gleich wichtig für den Kunden sind. In der Analyse der langen Strecke war die Variable Reisezeit geringfügig wichtiger als die Kosten. Durch die Veränderung der einzelnen Teilnutzenwerte, kann auch eine Veränderung des Gesamtnutzens für die beiden Varianten festgestellt werden.

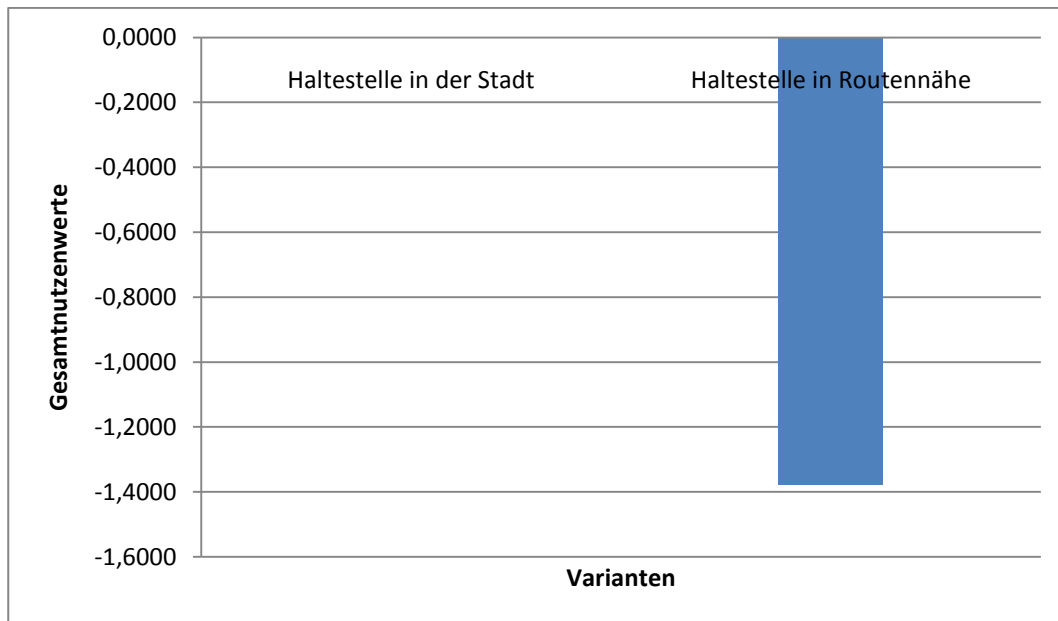


Abbildung 11: Gesamtnutzenwerte der Varianten für kurze Strecken (e.D.)

In Abbildung 11 kann festgestellt werden, dass der Gesamtnutzen für die Haltestellen in Routennähe bei – 1,3785 liegt. Damit ist dieser deutlich niedriger als der Nutzen für die Variante 1. So zeigt sich, dass für kurze Strecken der Nutzen von Haltestellen im Innenstadtbereich für den Kunden höher ist, als für Haltestellen in Routennähe. Dies liegt daran, dass die Anreise- und Abreisezeiten zu und von den Fernbushaltestellen in der Innenstadt kürzer sein können, so wie in dieser Schätzung angenommen.

Güteprüfung	
LL-Wert	- 239,5214
LL0-Wert	- 330,6312
LLR-Wert	- 182,2197
p-Wert (Chi²-Verteilung)	0,0000
McFadden-R²	27,56%
Trefferquote	80,04%

Tabelle 19: Güteprüfung des Modells (kurze Strecken)

Nach der Ermittlung der Gesamtnutzenwerte für beide Varianten soll das Modell auf die globale Güte geprüft werden (siehe Tabelle 19). Dafür werden wiederum der Likelihood-Ratio-

Test und die Trefferquote hinzugezogen. Es zeigt sich, dass der *LL*-Wert für das Modell der kurzen Strecken um einiges höher liegt als der *LL*-Wert der langen Strecken. Der LogLikelihood-Wert des Nullmodells ist dafür sehr ähnlich zu dem vorangegangenen. Des Weiteren liegt auch der p-Wert für den Likelihood-Ratio-Test wieder bei 0,0000 und sagt somit aus, dass das Modell statistisch hochsignifikant ist. Allerdings liegt der McFadden-R²-Wert hier nur bei 27,56 Prozent. Dieser Wert ist nach allgemeinen Angaben zwar immer noch zufriedenstellend, liegt aber 36,1 Prozent unter dem Wert für die langen Strecken. Auch die Trefferquote ist für diese Analyse etwas schlechter als vorher, aber mit 80,04 Prozent immer noch sehr hoch. Das Modell für die kurzen Strecken besitzt somit einen befriedigenden Modellfit, ist hochsignifikant und kann mit hoher Wahrscheinlichkeit die Wahl der Probanden vorhersagen. Neben der globalen Prüfung können hier auch die einzelnen Koeffizienten betrachtet werden.

Ausprägung	Hohe Kosten	Lange Reisezeit
Schätzwert (b_{jm})	- 0,6892	- 0,6892
LL0 _j -Wert (reduziert)	- 260,1969	- 260,1969
LL-Wert (vollständig)	- 239,5214	- 239,5214
Differenz	20,6755	20,6755
LLR _j	41,3510	41,3510
p-Wert	0,0000	0,0000
Signifikanz	Stark signifikant	Stark signifikant

Tabelle 20: Prüfung der Koeffizienten der kurzen Strecken

Tabelle 20 zeigt die Ergebnisse der Prüfung der einzelnen Koeffizienten für das Modell mit kurzen Strecken. Wie auch in Kapitel 4.3.1 wird der Likelihood-Ratio-Test angewendet. Auch hier zeigt sich eine sehr starke Signifikanz der Variablen für das Modell. Des Weiteren ist zu erkennen, dass beide Variablen genau den gleichen Einfluss auf das Modell haben und somit gleich bedeutend für die Entscheidung der Probanden sind.

Wie in Kapitel 4.3.1 ist die Analyse mit einer weiteren Aufspaltung der Reisezeiten wiederholt worden. Anders als bei den langen Strecken ist jedoch die Reisezeit im Bus für beide

Situationen gleich, da keine zusätzlichen Haltestellen angefahren werden. So können in diesem Fall nur drei Variablen mit jeweils zwei Ausprägungen untersucht werden.

Eigenschaft	Ausprägung	Teilnutzen	rel. Wichtigkeit
Kosten	Hoch	- 0,4595	33,33 %
	Niedrig	0,0000	
Zeit (bis Haltestelle)	Lang	- 0,4595	33,33 %
	Kurz	0,0000	
Zeit (von Haltestelle)	Lang	- 0,4595	33,33 %
	Kurz	0,0000	

Tabelle 21: Teilnutzenwerte der kurzen Strecke mit Aufspaltung der Reisezeit

In Tabelle 21 zeigt sich, dass, wie zuvor, die hohen Kosten und langen Reisezeiten einen negativen Teilnutzenwert besitzen. Ebenso ist die Wichtigkeit der Eigenschaften immer noch gleich verteilt. Dies hat zur Folge, dass die Situation 1 mit Haltestellen im Zentrum einen höheren Gesamtnutzenwert besitzt als Situation 2. Diese sind in Abbildung 12 dargestellt.

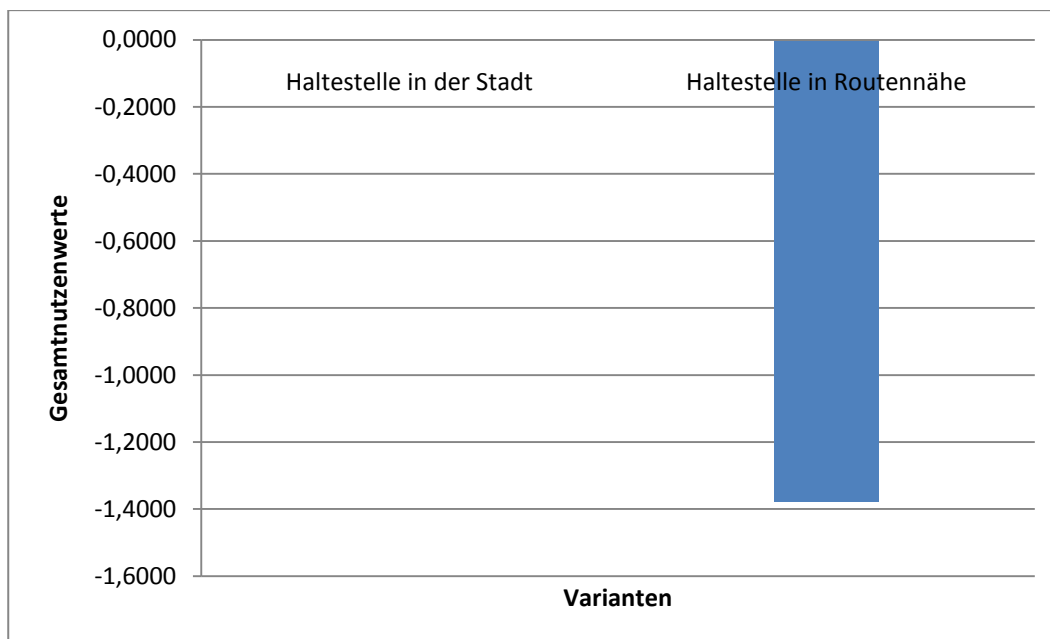


Abbildung 12: Gesamtnutzenwerte der kurzen Strecke mit Aufspaltung der Reisezeit (e.D.)

So können auch hier die Ergebnisse aus der vorangegangenen Conjoint-Analyse bestätigt werden.

Mit Hilfe der Formel 4.12 können auch für die kurzen Strecken die Wahrscheinlichkeiten für die Wahl der Probanden errechnet werden.

$$prob(1|2) = \frac{e^{0,0000}}{e^{0,0000} + e^{-1,3785}} = 0,7987$$

$$prob(2|1) = 1 - 0,7987 = 0,2013$$

Es ist zu erkennen, dass bei den kurzen Strecken die Situation 1 mit der Haltestelle im Stadtzentrum von den Probanden bevorzugt werden würde, wenn beide Angebote gleichzeitig zur Verfügung stehen. Auch hier wird die Fahrt mit der kürzeren Reisezeit bevorzugt. Wird der Rationalitätsparameter $\rho = 0,5$ mit einbezogen, sehen die Ergebnisse wie folgt aus:

$$prob(1|2) = 0,6658$$

$$prob(2|1) = 0,3342$$

Hier wird, genau wie bei den langen Strecken, eine Anpassung zu Gunsten der schwächeren Alternative vorgenommen, um den realen Marktverhältnissen näher zu kommen. Dennoch wählen knapp zwei Drittel die Alternative 1. Auch hier ist eine Disaggregation der Nutzenwerte nicht sinnvoll, da eine weitestgehend homogene Nutzenvorstellung der Probanden vorliegen sollte.

4.3.3 Lange und kurzen Strecken zusammen

Nach der getrennten Betrachtung der kurzen und langen Strecken, sollen diese in einer weiteren Analyse zusammen untersucht werden. Hier soll vor allem die Wichtigkeit der Variablen Zeit und Kosten genauer analysiert werden. Dabei sind die langen und kurzen Strecken genau wie vorher definiert. Allerdings werden pro Person nun acht Auswahlentscheidungen betrachtet, sodass insgesamt 952 Entscheidungen in die Analyse mit einfließen.

	Hohe Kosten (b ₁₁)	Niedrige Kosten (b ₁₂)	Lange Reisezeit (b ₂₁)	Kurze Reisezeit (b ₂₂)
Absolute Häufigkeit	539	414	129	824
Auswahlsituationen	952	952	952	952
Anteil	0,5662	0,4349	0,1355	0,8665
Startwert	0,1313	0,0000	- 0,7300	0,0000
Schätzwerte	0,6093	0,0000	- 1,9878	0,0000

Tabelle 22: Ermittlung Startwerte und Schätzwerte nach ML-Schätzung beider Strecken

Es zeigt sich, dass die Richtungen der Schätzwerte ähnlich zu den Ergebnissen der langen Strecken sind (siehe Tabelle 22). Allerdings kann festgestellt werden, dass die lange Reisezeit im Vergleich zu der kurzen Reisezeit einen stark negativen Nutzen aufweist. Die Auswahlwahrscheinlichkeiten für beide Alternativen variieren je nach Länge der Strecke.

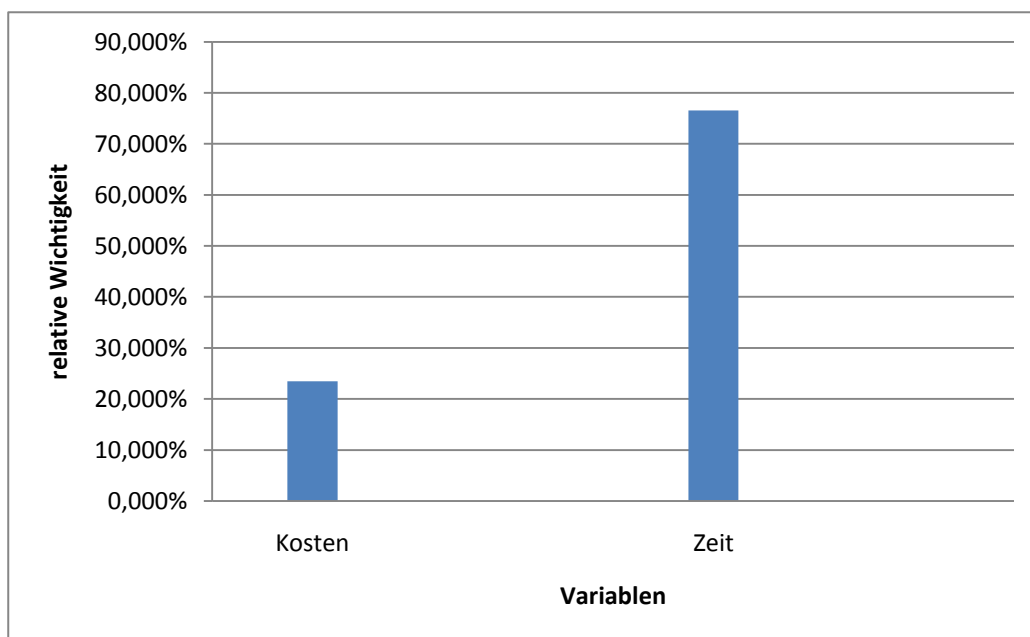


Abbildung 13: Relative Wichtigkeit der Variablen bei beiden Strecken (e.D.)

Im Anschluss an die Berechnung der Teilnutzen der einzelnen Merkmalsausprägungen, kann die relative Wichtigkeit der Variablen bestimmt werden, welche in Abbildung 13 dargestellt ist. Es zeigt sich, anders als in den vorangegangenen Analysen, dass die Reisezeit (76,54 Prozent) viel wichtiger für die Entscheidung eines Kunden ist, als die Reisekosten (23,46 Prozent). So fällt der geringe Preisunterschied von drei Euro bei beiden Alternativen nur wenig

ins Gewicht und es kann festgehalten werden, dass Kunden zuerst auf die Zeit schauen und nach möglichst kurzen Reisezeiten suchen.

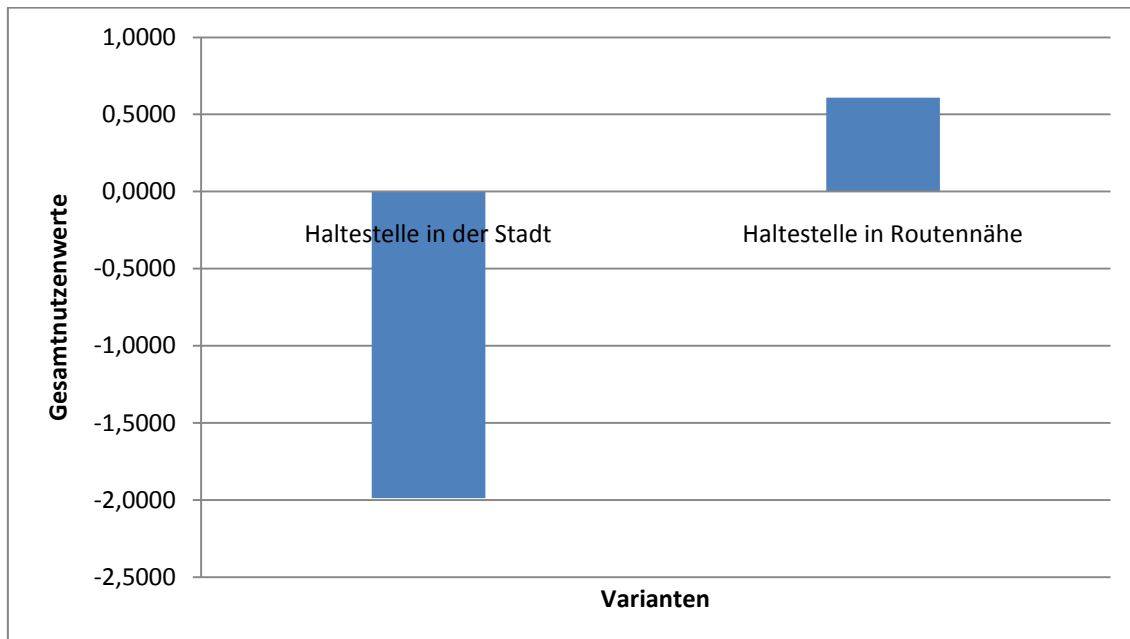


Abbildung 14: Gesamtnutzenwerte der Varianten für beide Strecken (e.D.)

In Abbildung 14 sind die resultierenden Gesamtnutzenwerte für beide Varianten abgebildet. Auch wenn lange und kurze Strecken in einer Analyse gemeinsam betrachtet werden, zeigt sich, dass die Haltestellen in Routennähe einen deutlich höheren Nutzen für die Kunden besitzen, als die Haltestellen in der Stadt. Dies kann daran liegen, dass bei langen Strecken sehr viel Zeit eingespart wird, wenn die Haltestellen in Routennähe liegen. Gleichzeitig ist der Zeitverlust für die kurzen Strecken gering, da sich die Reisezeiten im ÖPNV in der Startstadt und Zielstadt nur geringfügig voneinander unterscheiden.

Tabelle 23 zeigt die wichtigen Ergebnisse zur Prüfung der Güte des Gesamtmodells. Es zeigt sich, dass der LL-Wert und der LL_0 -Wert sehr viel größer sind, als bei den vorangegangenen Modellen. Das McFadden- R^2 liegt genau zwischen den Pseudo- R^2 -Werten der langen (63,66 Prozent) und kurzen (27,56 Prozent) Strecken. Da ein McFadden- R^2 zwischen 20 und 40 Prozent als gute Anpassung angesehen werden kann, ist die Güte des Modells sehr zufriedenstellend. Auch die Trefferquote für beide Strecken liegt ebenfalls genau zwischen den Trefferquoten der beiden anderen Modelle und ist mit 86,55 Prozent sehr gut. So können fast 90 Prozent

der Entscheidungen von Kunden richtig vorhergesagt werden. Durch den p-Wert von 0,0000 kann festgestellt werden, dass das Modell statistisch hochsignifikant ist.

Güteprüfung	
LL-Wert	- 359,4241
LL0-Wert	- 660,5693
LLR-Wert	- 602,2903
p-Wert (Chi²-Verteilung)	0,0000
McFadden-R²	45,59%
Trefferquote	86,55%

Tabelle 23: Güteprüfung des Modells (lange und kurze Strecken)

Nach der Güteprüfung des Gesamtmodells können wieder die einzelnen Variablen auf ihre Signifikanz hin untersucht werden. Es zeigt sich, dass beide Merkmalsausprägungen für das Modell von sehr großer Bedeutung sind. Des Weiteren kann festgestellt werden, dass die Differenz zwischen dem vollständigen LL-Wert und dem reduzierten LL_{0j}-Wert für die lange Reisezeit bei 306,6214 liegt. Daraus lässt sich schließen, dass die Reisezeit mehr Einfluss auf das Modell hat, wie auch die relative Wichtigkeit der einzelnen Variablen bestätigt. Werden also kurze und lange Strecken zusammen betrachtet, ist die Reisezeit für den Kunden das ausschlaggebende Merkmal, welches zu einer Entscheidung führt.

Ausprägung	Hohe Kosten	Lange Reisezeit
Schätzwert (b_{jm})	0,6093	- 1,9878
LL0_j-Wert (reduziert)	- 378,7814	- 666,0456
LL-Wert (vollständig)	- 359,4241	- 359,4241
Differenz	19,3573	306,6214
LLR_j	38,7145	613,2429
p-Wert	0,0000	0,0000
Signifikanz	Stark signifikant	Stark signifikant

Tabelle 24: Prüfung der Koeffizienten für beide Strecken

Auch hier können mit der Formel 4.12 die Wahrscheinlichkeiten für die Wahl der Probanden errechnet werden.

$$prob(1|2) = \frac{e^{-1,9878}}{e^{-1,9878} + e^{0,6093}} = 0,0693$$

$$prob(2|1) = 1 - 0,0693 = 0,9307$$

Es ist zu erkennen, dass für eine gemeinsame Betrachtung der langen und kurzen Strecken die Variante 2 von den Fernbuskunden bevorzugt wird, wenn beide Angebote gleichzeitig zur Verfügung stehen. Auch hier wird die Fahrt mit der kürzeren Reisezeit bevorzugt. Wird der Rationalitätsparameter $\rho = 0,5$ mit einbezogen, sehen die Ergebnisse wie folgt aus:

$$prob(1|2) = 0,2144$$

$$prob(2|1) = 0,7856$$

Auch hier wird, wie zuvor, eine Anpassung zu Gunsten der schwächeren Alternative vorgenommen. Trotzdem wählen fast 80 Prozent der Probanden die Alternative 2. Dies zeigt, dass bei der gemeinsamen Betrachtung von langen und kurzen Strecken die Reisezeitersparnis bei Haltestellen in Routennähe einen großen Einfluss auf die Entscheidung hat.

Wie in Kapitel 4.3.1 und 4.3.2 wurde die Analyse mit einer binären Kodierung der Merkmalsvariablen durchgeführt. Das bedeutet, dass für die Reisezeiten und Reisekosten nur eine 1 oder 0 angegeben worden sind, je nachdem um welche Variante mit welchen Ausprägungen es sich handelt. Allerdings können auch die tatsächlichen Reisekosten und Reisezeiten für jede Strecke in die Analyse mit eingezogen werden. Hier zeigt sich ebenfalls, dass die Reisezeit das wichtigste Entscheidungsmerkmal darstellt.

Eigenschaften	Ausprägung	Teilnutzen	zentriert	Range	Rel. Wichtigkeit
Kosten	Hoch	- 0,0285	- 0,0143	0,0285	8,00 %
	Niedrig	0,0000	0,0143		
Reisezeit	Lang	- 0,3279	- 0,1640	0,3279	92,00 %
	kurz	0,0000	0,1640		

Tabelle 25: Rel. Wichtigkeit der Variablen beider Strecken mit expliziten Werten

Tabelle 25 zeigt die Teilnutzenwerte und die relative Wichtigkeit der Variablen. Es kann festgestellt werden, dass sowohl hohe Kosten als auch eine lange Reisezeit einen negativen Teilnutzenwert besitzen. Allerdings ist der Teilnutzen für eine lange Reisezeit mit $-0,3279$ gering. Ist die Reisezeit sehr lang, sinkt der Nutzen für den Kunden sehr stark ab. So sind auch für beide Varianten die Gesamtnutzenwerte negativ, aber für die Haltestellen in Routennähe liegt dieser bei $-0,0285$ NE und für die Haltestellen innerhalb der Stadt bei $-0,3279$ NE (siehe Abbildung 15).

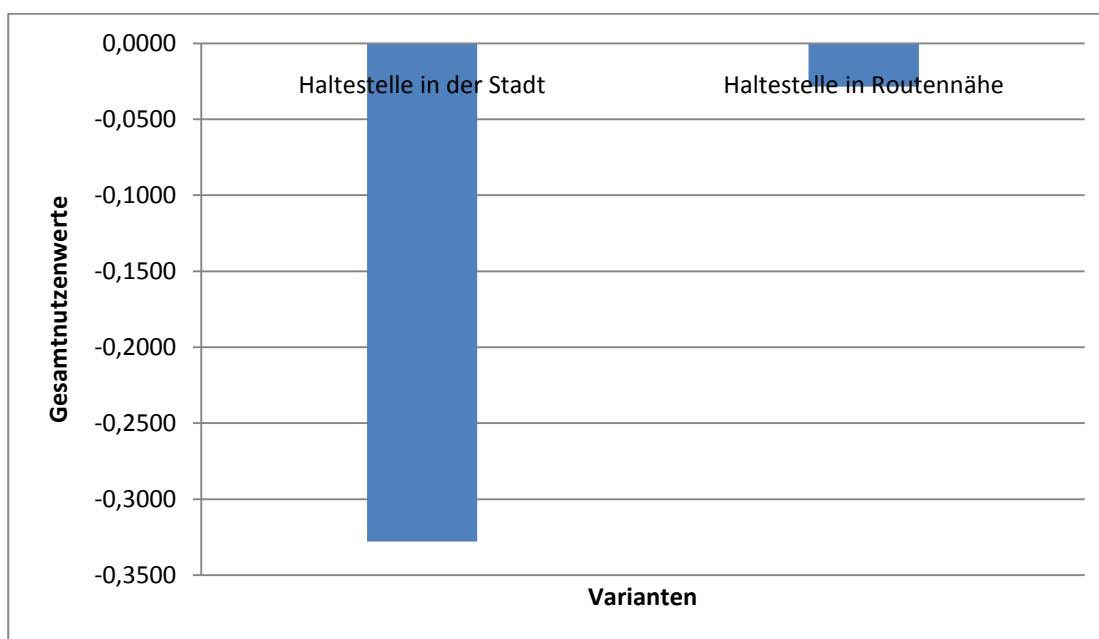


Abbildung 15: Gesamtnutzenwerte der Varianten für beide Strecken mit expliziten Werten (e.D.)

Auch hier sind beide Merkmalsausprägungen sehr signifikant für das Modell, aber ähnlich wie bei der binär kodierten Analyse, ist die Reisezeit von größerer Bedeutung. So zeigt sich, dass die Haltestellen in Routennähe für alle Kunden, unabhängig von der Länge der Strecke, einen größeren Nutzen besitzen.

Eine weitere Erkenntnis, welche durch das Verwenden der genauen Reisezeiten und Kosten gewonnen werden kann, soll ein Zeitwert sein, der aussagt, ab welcher Reisezeit die eine Variante besser als die Andere ist und wann die Kosten nicht mehr ausschlaggebend für die Entscheidung ist. Allerdings ist dabei anzumerken, dass für verschiedene Strecken kein genereller Wert existieren kann. Dabei soll ein Beispiel auf diese Problematik aufmerksam machen. Eine Strecke von drei Stunden Reisezeit in Brandenburg ist nicht mit einer Strecke und der glei-

chen Reisezeit in Nordrhein-Westfalen zu vergleichen. Dies liegt vor allem an der Anzahl der Zwischenstopps, welche auf den jeweiligen Strecken angefahren werden. Die Fahrt in Brandenburg, welche drei Stunden dauert, hat mit hoher Wahrscheinlichkeit wenige bis gar keine Zwischenhalte, da dort die Bevölkerungsdichte gering ist. Die Fahrt in Nordrhein-Westfalen, welche drei Stunden dauert, hat durch die hohe Bevölkerungsdichte und sehr vielen Großstädte eine größere Anzahl an Haltestellen, die angefahren werden können. So ist nun bei beiden Strecken die Reisezeit gleich, aber die gefahrene Strecke ist unterschiedlich. Bei wenigen bis gar keinen Zwischenhalten ist, wie in Kapitel 4.3.2, eine Haltestelle im Stadtzentrum von größerem Nutzen für die Kunden. Werden mehrere Städte angefahren, wie in Kapitel 4.3.1, besitzt die Haltestelle in Routennähe einen höheren Nutzen. Daher zeigt sich, dass es keine pauschale Reisezeit geben kann, an der die eine Variante besser als die andere ist. Es kann höchstens die Anzahl der Zwischenstopps und die damit verbundene Zeitdifferenz näher untersucht werden. Dabei kann grob festgestellt werden, dass ab ca. zwei zusätzlich angefahrenen Haltestellen, die Variante mit Haltestellen in Routennähe einen größeren Nutzen für die Kunden besitzt, als die Haltestellen im Stadtzentrum. Dies liegt daran, dass sich dann die Zeitersparnis für Zwischenhalte bei Haltestellen in Routennähe und die zusätzliche Reisezeit zu und von den Haltestellen weitestgehend ausgleichen. Dabei kommt es auch auf den genauen Start- bzw. Zielort an, der in dieser Arbeit nicht festgestellt worden ist. Bei einem bzw. keinem Zwischenstopp besitzen die Haltestellen im Stadtzentrum einen höheren Nutzen.

4.3.4 Ergebnis der auswahlbasierten Conjoint-Analyse

Werden nun die Ergebnisse der Kapitel 4.3.1, 4.3.2 und 4.3.3 betrachtet, lässt sich ein großer Unterschied feststellen. Es zeigt sich, dass bei den langen Strecken die Reisezeit marginal wichtiger ist, als die Reisekosten. Für die kurzen Strecken kann für beide Variablen die gleiche Wichtigkeit festgestellt werden. Des Weiteren nehmen die Probanden bei langen Strecken einen höheren Preis in Kauf, wenn sie dadurch eine Zeitersparnis erreichen. So kann gezeigt werden, dass für Strecken von Fernbussen, die mehrere Städte anfahren, eine Haltestelle in Routennähe einen höheren Nutzen besitzt, als Haltestellen im Stadtzentrum. Allerdings muss dabei eine gute ÖPNV-Anbindung gewährleistet sein, um die An- und Abreisezeiten nicht zu lang werden zu lassen. Für kurze Strecken, die keinen Zwischenstopp einlegen, sind die Haltestellen im Stadtzentrum besser geeignet. Das liegt daran, dass es auf der Strecke keine Zeitersparnis gibt und die An- und Abreisezeiten bei Haltestellen in Routennähe höher sind, als

die bei den Haltestellen im Zentrum. Auch wenn sowohl lange als auch kurze Strecken gemeinsam betrachtet werden, besitzen die Haltestellen in Routennähe einen größeren Nutzen für die Kunden. Die wichtigste Rolle spielt dabei wieder die enorme Zeitersparnis. Es zeigt sich, dass vor allem die Reisezeit Einfluss auf die Auswahlentscheidung der Probanden hat. Diese nehmen einen höheren Preis in Kauf, wenn sich die Reisezeit entsprechend verringert. So kann eventuell von Seiten der Fernbusunternehmen über weitere Haltestellen in Routennähe mit ÖPNV-Anbindung auf den Strecken nachgedacht werden, auf denen viele andere Städte zusätzlich angefahren werden.

4.4 Logistische Regression

Im Folgenden Kapitel sollen jeweils für die langen und kurzen Strecken logistische Regressionen durchgeführt werden, um den Einfluss der sozioökonomischen Merkmale auf die Auswahlentscheidungen zu untersuchen. Dabei wird folgende binäre abhängige Variable y verwendet:

$$Y = \begin{cases} = 1, & \text{wenn die Variante 2 (Haltestelle in Routennähe) gewählt wird} \\ = 0, & \text{wenn die Variante 1 (Haltestelle im Stadtzentrum) gewählt wird} \end{cases}$$

Es stehen folgende sozioökonomische Variablen zur Verfügung: Geschlecht, Alter, Beruf, Einkommen, Schulabschluss, Berufsausbildung und die PKW-Verfügbarkeit. Da die Variablen Beruf, Schulabschluss und Berufsausbildung nach dem Rangkorrelationskoeffizienten nach Spearman (Formel 4.8) eine hohe Abhängigkeit aufweisen, wird nur das Merkmal Beruf in der logistischen Regression berücksichtigt. Da zwischen dem Alter und dem Beruf ebenfalls eine starke Korrelation zu erkennen ist, muss hier entschieden werden, welche Variable in der Untersuchung nicht berücksichtigt wird. Mit Hilfe des Akaike Informations Criteria (AIC) werden nun zwei Modelle miteinander verglichen. Dabei soll herausgefunden werden, ob die Variable Alter oder das Merkmal Berufsstatus in die Regression mit einbezogen werden sollen. Das AIC vergleicht die Qualität eines Modells in Relation zu einem Anderen (Akaike, H. (1973)). Es kann zwar ein besserer Wert für das Modell mit der Variable Alter berechnet werden (219,86) als für das Modell mit dem Merkmal Berufsstatus (228,20), hier wird aber dennoch die Regression mit dem Merkmal Berufsstatus verwendet, da die Parameterwerte und Standardfehler sehr viel geringer ausfallen. Das Merkmal Beruf wird in die Klassen „Ausbildung“, „Erwerbstätig“ und „nicht Erwerbstätig“ eingeteilt, wobei „Ausbil-

„ung“ als Referenzkategorie verwendet wird. Für das Einkommen wird jeweils die Klassenmitte als Repräsentant der Klasse gewählt. So kann für die langen und kurzen Strecken die folgende Regressionsgleichung betrachtet werden:

$$Y = a_0 + a_1 * \text{Geschlecht} + a_2 * \text{Erwerbstätig} + a_3 * \text{nicht Erwerbstätig} + a_4 * \text{Einkommen} + a_5 * \text{PKW} \quad (4.13)$$

4.4.1 Lange Strecken

Die Probanden sollten für die langen Strecken jeweils vier Auswahlentscheidungen treffen, sodass insgesamt 476 Entscheidungen in der Regression betrachtet werden können.

Variable	
Intercept	2,1945*** (0,4029)
Geschlecht	0,6929* (0,4112)
Erwerbstätig	-1,2866* (0,6894)
Nicht Erwerbstätig	0,2547 (1,0551)
Einkommen	0,0010* (0,0005)
PKW-Verfügbarkeit	-0,8996** (0,4158)
McFadden-R²	0,0784
Nagelkerke-R²	0,0970
Anzahl Beobachtungen	476

Signifikanzniveaus werden wie folgt ausgewiesen: * = 90%, ** = 95%, *** = > 99%

Tabelle 26: Ergebnisse logistische Regression der langen Strecken

In Tabelle 26 sind sowohl die geschätzten Koeffizienten für alle Merkmale mit Signifikanzniveau und Standardfehler als auch der McFadden- und Nagelkerke-R² ausgewiesen. Es zeigt sich, dass alle Variablen außer „nicht Erwerbstätig“ signifikant für das Modell

sind. Eine genaue Interpretation der Parameter ist in diesem Zustand noch nicht möglich. Es kann lediglich die Richtung des Einflusses jedes Merkmal abgelesen werden. Dabei haben die Erwerbstätigkeit und die PKW-Verfügbarkeit einen negativen Einfluss auf die Wahl für Haltestellen in Routennähe. Da die Ausprägung „Ausbildung“ die Referenz für das Merkmal Berufsstatus ist, kann hier gesagt werden, dass erwerbstätige Personen im Gegensatz zu Personen in Ausbildung, einen negativen Einfluss auf die abhängige Variable besitzen. Alle anderen Variablen begünstigen die Situation mit Haltestellen in Routennähe. Dies zeigt, dass erwerbstätige, männliche Personen, die in Besitz eines eigenen PKW sind, sich eher für die andere Variante entscheiden. Mit Hilfe der Odds Ratio, welche auch als Effektkoeffizienten bezeichnet werden, wird eine Interpretation erleichtert. Diese sind in Tabelle 27 dargestellt.

Variable	Odds Ratio ($e^{a_{etg}}$)
Intercept	8,9752
Geschlecht	2,0000
Erwerbstätig	0,2762
Nicht Erwerbstätig	1,2901
Einkommen	1,0010
PKW-Verfügbarkeit	0,4067

Tabelle 27: Odds Ratio der langen Strecken

Die Odds Ratio sagen aus, wie stark der Einfluss auf das Wahrscheinlichkeitsverhältnis ist, wenn die Variable um eine Einheit erhöht wird. Dabei kann davon ausgegangen werden, dass ein Wert der Odds Ratio größer eins die Auswahlwahrscheinlichkeit für $y = 1$ erhöht, während ein Wert kleiner eins diese Wahrscheinlichkeit sinken lässt (Backhaus (2011), S. 265/266). So zeigt sich, dass Geschlecht und nicht Erwerbstätigkeit einen positiven Einfluss auf die Variante 2 besitzen, während Erwerbstätigkeit und PKW-Verfügbarkeit eher die andere Variante begünstigen. Die Variable Einkommen hat nur einen geringen positiven Einfluss auf $y = 1$. Es muss berücksichtigt werden, dass es sich bei fast allen Variablen um binäre oder nominalskalierte Merkmale handelt und so die Einflüsse kleiner ausfallen können als berechnet.

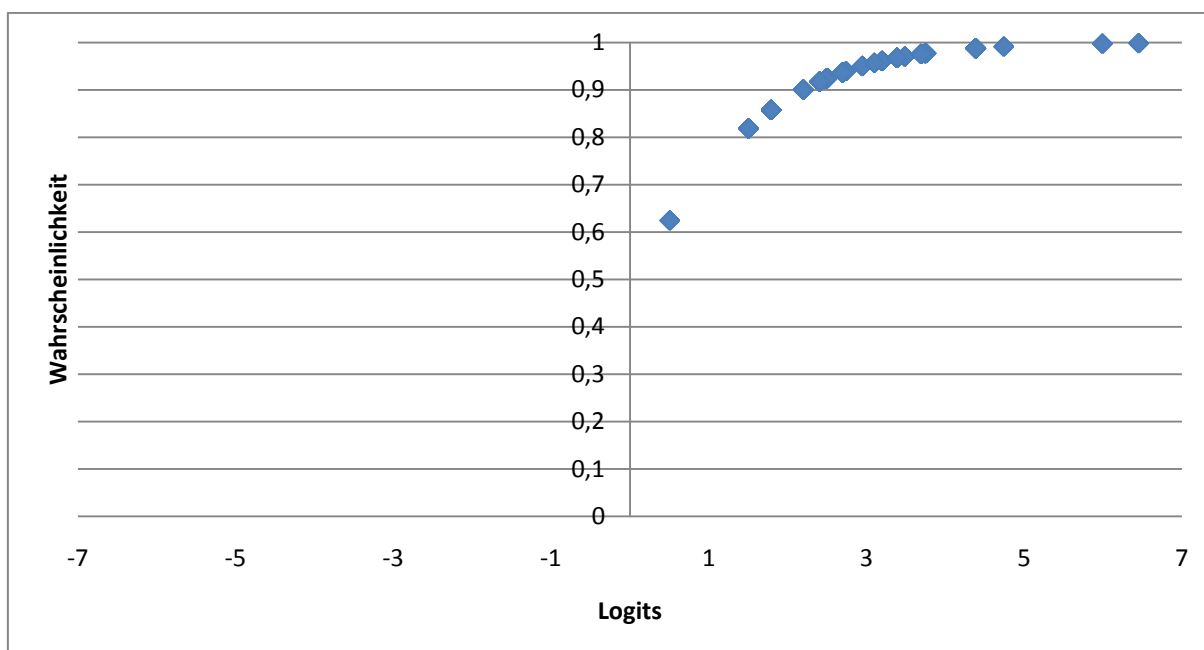


Abbildung 16: Wahrscheinlichkeitsverteilung lange Strecken (e.D.)

In Abbildung 16 ist die Wahrscheinlichkeitsverteilung für die langen Strecken abgebildet. Alle Entscheidungen mit einem $prob_g$ größer 0,5 werden der Variante 2, Haltestellen in Routennähe, zugeteilt. Hier zeigt sich, dass alle Entscheidungen bei langen Strecken der Variante 2 zugeordnet werden. Dieses Ergebnis kann auch mittels Klassifikationstabelle dargestellt werden, welche das Gütemaß des gesamten Modells beurteilen kann.

Von/nach	y = 1	y = 0	Summe
y = 1	444	0	444
y = 0	32	0	32
Summe	476	0	476

Trefferquote y = 1 100,00 %

Trefferquote y = 0 0,00 %

Tabelle 28: Klassifikationstabelle lange Strecken

Neben der Klassifikationstabelle können der Likelihood-Ratio-Test, das McFadden- R^2 und das Nagelkerke- R^2 Informationen über die gesamte Anpassung des Modells geben (siehe Tabelle 28). Sowohl der McFadden- R^2 -Wert als auch der Nagelkerke- R^2 -Wert können in Tabelle 26 abgelesen werden. Beide sind mit unter zehn Prozent sehr niedrig und weisen nicht unbedingt auf einen guten Modellfit hin. Dahingegen kann durch den Likelihood-Ratio-Test festgestellt werden, dass das aufgestellte Modell für die betrachteten Entscheidungen signifikant ist.

Im Zuge der Ausreißerdiagnostik werden die Residuenwerte der einzelnen Auswahlentscheidungen näher betrachtet (siehe Abbildung 17)

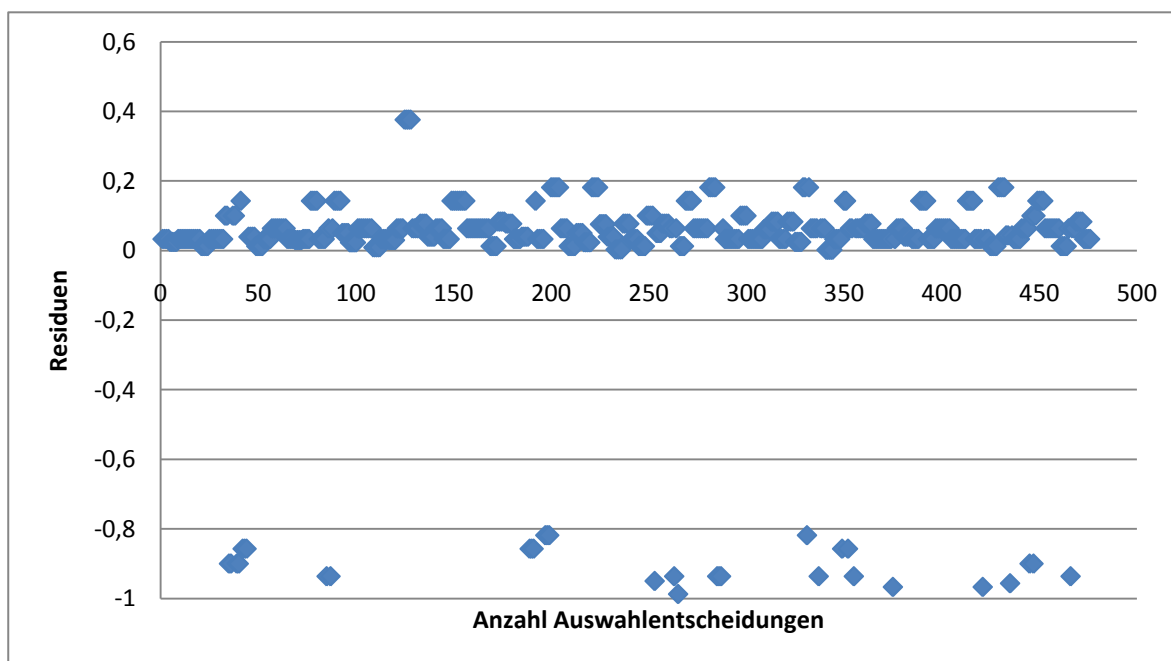


Abbildung 17: Ausreißer lange Strecken (e.D.)

Dabei können Werte größer 0,5 oder kleiner -0,5 als Ausreißer angenommen werden. Es ist zu erkennen, dass hier nur Ausreißer vorkommen, die kleinere Residuenwerte als -0,5 besitzen. Hierbei handelt es sich um alle Entscheidungen, bei denen die Variante 1 mit Haltestellen im Stadtzentrum von den Probanden gewählt worden ist.

Zuletzt können die einzelnen Merkmalsvariablen auf ihre Signifikanz für das Modell hin untersucht werden. Es wird die Wald-Statistik verwendet, welche die Nullhypothese testet, dass die Variable keinen Einfluss auf das Modell besitzt und somit gleich Null ist. Dabei kann für alle Variablen, außer für die Variable Berufsstatus (Erwerbstätig und nicht Erwerbstätig), eine Signifikanz für das Modell festgestellt werden.

4.4.2 Kurze Strecken

Wie auch bei den langen Strecken, konnten die Probanden bei den kurzen Strecken vier Auswahlentscheidungen treffen und somit werden auch hier 476 Entscheidungen in der Regression betrachtet.

Variable	
Intercept	-1,4336*** (0,1997)
Geschlecht	-0,7660*** (0,2545)
Erwerbstätig	0,8563** (0,3794)
Nicht Erwerbstätig	1,9437*** (0,4475)
Einkommen	0,0002 (0,0002)
PKW-Verfügbarkeit	-0,4067 (0,3012)
McFadden-R²	0,0821
Nagelkerke-R²	0,1250
Anzahl Beobachtungen	476

Signifikanzniveaus werden wie folgt ausgewiesen: * = 90%, ** = 95%, *** = > 99%

Tabelle 29: Ergebnisse logistische Regression der kurzen Strecken

Tabelle 29 zeigt die geschätzten Parameterwerte mit Signifikanzniveau und Standardfehler für die logistische Regression der kurzen Strecken. Es zeigt sich, dass Geschlecht, Erwerbstätigkeit und nicht Erwerbstätigkeit signifikant für das Modell sind. Dabei besitzt das Geschlecht ein negatives Vorzeichen und hat somit einen negativen Einfluss auf die Entscheidung für Variante 2. Erwerbstätigkeit und nicht Erwerbstätigkeit haben hingegen einen positiven Einfluss für die Haltestellen in Routennähe. Auch hier müssen für eine genauere Interpretation die Wahrscheinlichkeitsverhältnisse und die Odds Ratio betrachtet werden (siehe Tabelle 30).

Variable	Odds Ratio (e^{eig})
Intercept	0,2384
Geschlecht	0,4649
Erwerbstätig	2,3545
Nicht Erwerbstätig	6,9848
Einkommen	1,0002
PKW-Verfügbarkeit	0,6658

Tabelle 30: Odds Ratio der kurzen Strecken

Es zeigt sich, dass Erwerbstätigkeit und nicht Erwerbstätigkeit einen starken positiven Einfluss auf die Entscheidung für $y = 1$ haben. Das Einkommen hat nur einen sehr geringen positiven Einfluss auf die Variante 2. Geschlecht und PKW-Verfügbarkeit hingegen begünstigen eher die Wahl für die Haltestellen im Stadtzentrum. Aber auch hier muss berücksichtigt werden, dass es sich bei fast allen Merkmalen um binäre oder nominal-skalierte Variablen handelt.

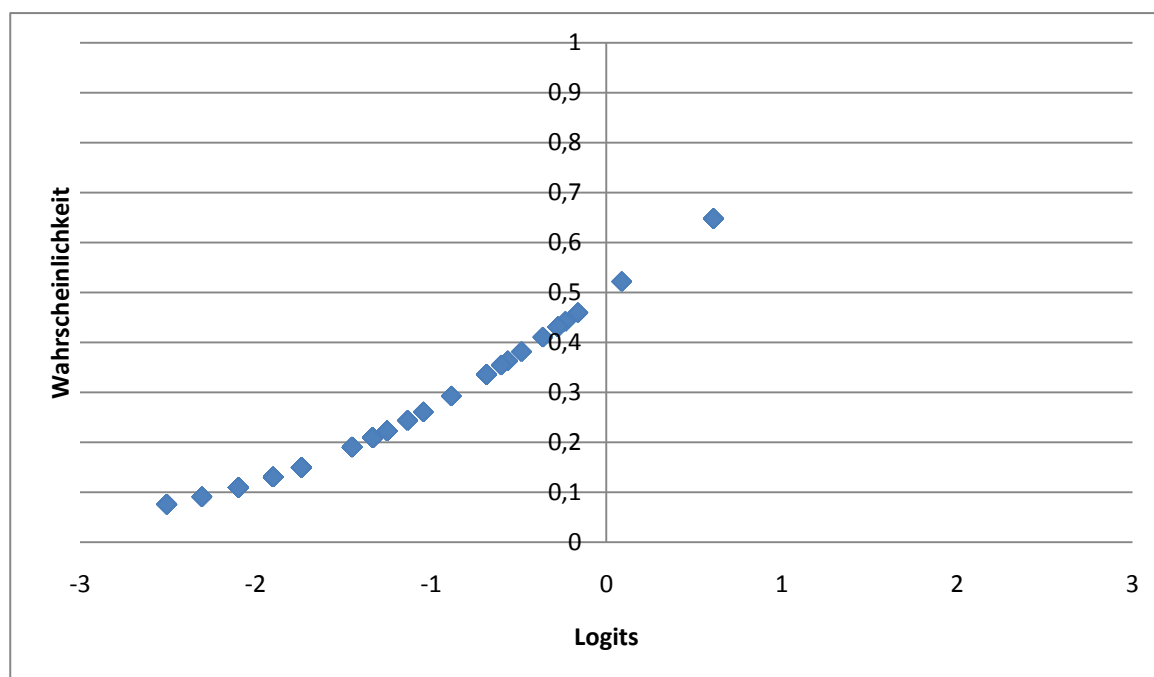


Abbildung 18: Wahrscheinlichkeitsverteilung kurze Strecken (e.D.)

Die Wahrscheinlichkeitsverteilung für die kurzen Strecken ist in Abbildung 18 dargestellt. Es zeigt sich, dass der Großteil der Entscheidungen eine Wahrscheinlichkeit von kleiner 0,5 be-

sitzen und somit der Variante 1 zugeordnet werden. Nur wenige Entscheidungen werden den Haltestellen in Routennähe zugeteilt. Auch dies kann mittels Klassifikationstabelle deutlicher dargestellt werden.

Von/nach	y = 1	y = 0	Summe
y = 1	7	91	98
y = 0	9	369	378
Summe	16	460	476

Trefferquote y = 1 7,14 %

Trefferquote y = 0 97,62 %

Tabelle 31: Klassifikationstabelle kurze Strecken

Es kann festgestellt werden, dass die meisten Entscheidungen die zugunsten der Variante 2 ausgefallen sind, falsch zugeordnet werden und die Trefferquote für y = 1 nur bei 7,14 Prozent liegt. Die Entscheidungen für Variante 1 werden in 97,62 Prozent der Fälle richtig zugeordnet (siehe Tabelle 31).

In Tabelle 29 sind wiederum die Pseudo-R²-Werte zu sehen. Der McFadden-R²-Wert liegt mit 0,0821 ein wenig höher als im Modell der langen Strecken. Auch der Nagelkerke-R²-Wert ist mit 0,1250 besser als im vorangegangenen Modell. Dennoch sind auch diese Werte nicht zufriedenstellend. Allerdings kann mittels Likelihood-Ratio-Test festgestellt werden, dass das aufgestellte Modell der kurzen Strecken für die untersuchten Entscheidungen signifikant ist.

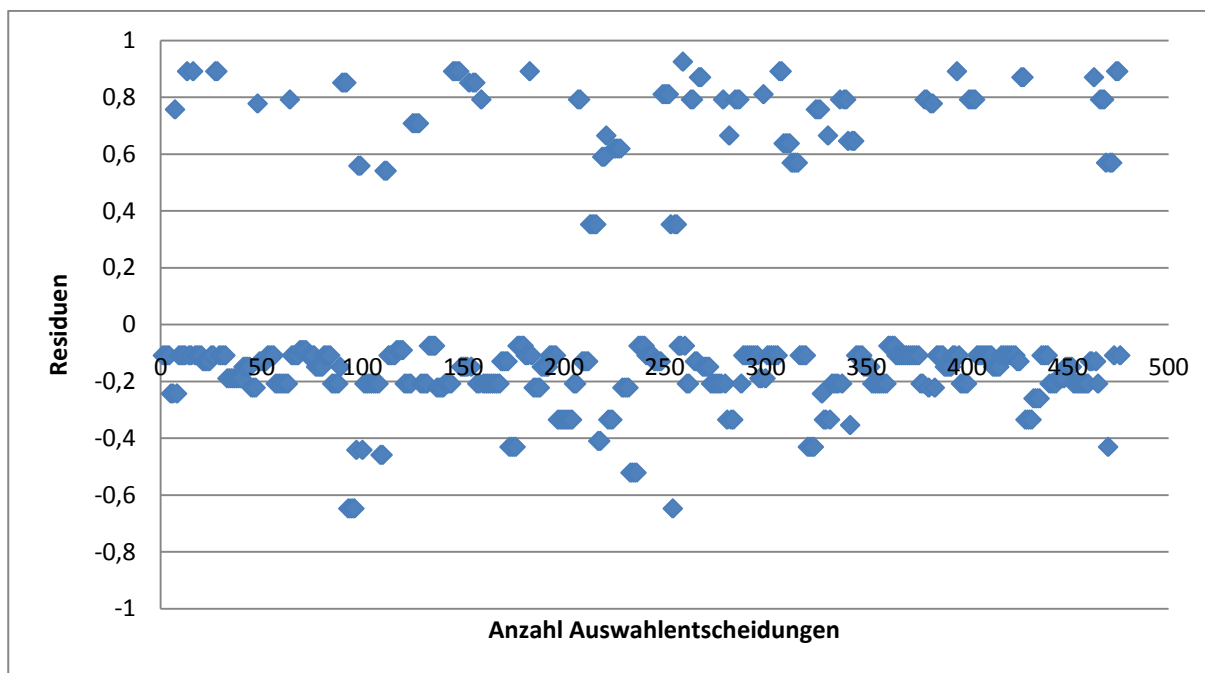


Abbildung 19: Ausreißer kurze Strecken (e.D.)

In Abbildung 19 sind die Residuenwerte für die Auswahlentscheidungen dargestellt, welche für die Ausreißerdiagnostik benötigt werden. Dabei zeigt sich, dass im Vergleich zu den langen Strecken mehr Ausreißer vorhanden sind. Allerdings sind hier die Ausreißer diejenigen Entscheidungen, die sich für die Variante 2 und gegen die Haltestellen im Stadtzentrum entschieden haben.

Anschließend können die einzelnen Variablen auf ihre Signifikanz untersucht werden. Anders als im Modell von Kapitel 4.4.1 sind alle Merkmale für das Modell signifikant und somit sind die Koeffizienten alle von Null verschieden.

4.4.3 Ergebnis der logistischen Regression

Nach der Analyse von langen und kurzen Strecken kann folgendes festgestellt werden. Für die Regression mit langen Strecken sind die Variablen Geschlecht, Erwerbstätigkeit, Einkommen und PKW-Verfügbarkeit signifikant. Es zeigt sich, dass alle Auswahlentscheidungen der abhängigen Variable $y = 1$ (Haltestelle in Routennähe) zugeordnet werden. Dabei können die Entscheidungen, die eigentlich zu $y = 0$ (Haltestelle in der Stadt) gehören, als Ausreißer für diese Regression deklariert werden. Des Weiteren kann festgestellt werden, dass sich erwerbstätige, männliche Personen mit einem eigenen PKW eher für die Variante mit Haltestellen im

Stadtzentrum entscheiden. So wählen Personen, die sich noch in der Ausbildung befinden und somit durch ein geringeres Einkommen meist keinen PKW zur Verfügung haben, die Variante mit Haltestellen in Routennähe, welche zwar etwas teurer ist, aber auch sehr viel mehr Zeit spart. Bei der Analyse der kurzen Strecken sind ebenfalls die Variablen Geschlecht und Erwerbstätigkeit signifikant. Hinzu kommt die Dummyvariable „Nicht Erwerbstätig“. Der Großteil der Auswahlentscheidungen wird der Situation $y = 0$ zugeordnet, obwohl es einen größeren Anteil an Entscheidungen für die andere Situation gibt. Hier zeigt sich, dass sich männliche Personen, die sich nicht mehr in der Ausbildung befinden, für die Variante 2 mit Haltestellen in Routennähe entscheiden.

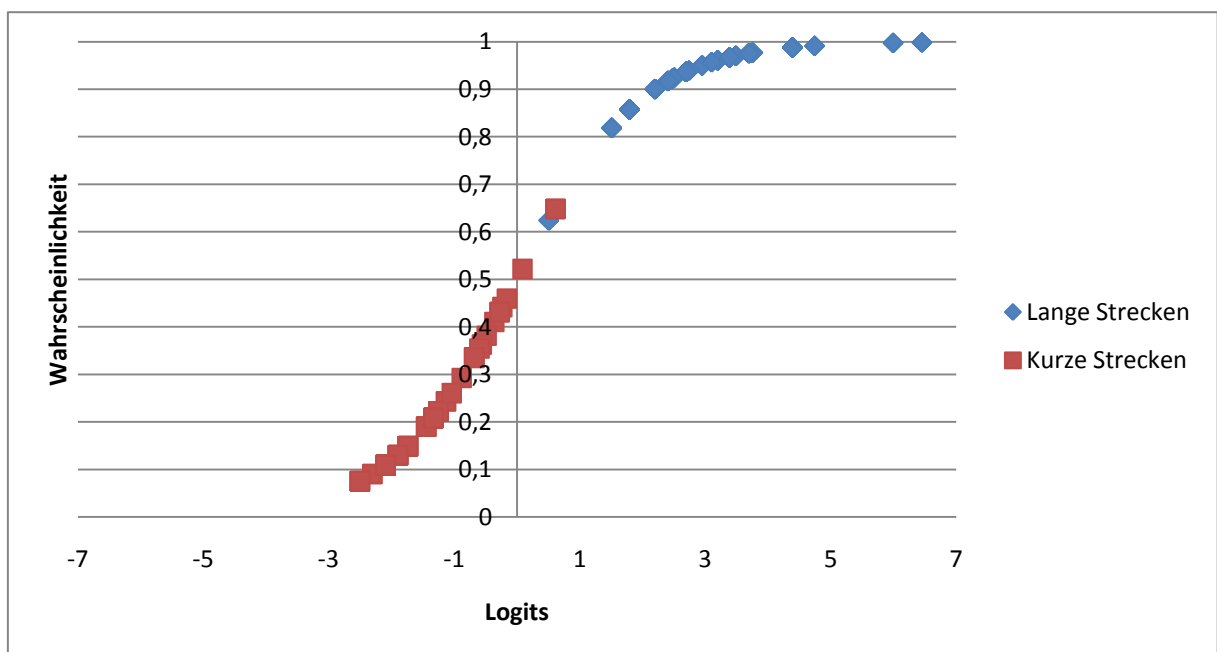


Abbildung 20: Wahrscheinlichkeitsverteilung für beide Strecken (e.D.)

Abbildung 20 zeigt die Wahrscheinlichkeitsverteilung sowohl für lange als auch für kurze Strecken. Es ist zu erkennen, dass die Auswahlentscheidungen bei langen Strecken zu der Variante 2 mit Haltestellen in Routennähe tendieren, während bei kurzen Strecken die Variante 1 mit Haltestellen in der Innenstadt bevorzugt werden.

4.5 Ergebnis der Analysen

Nun sollen die Ergebnisse aller Analysen noch einmal kurz zusammengefasst werden. Anhand der Zielgruppenanalyse (Kapitel 4.2) konnte der „typische“ Fernbusnutzer charakteri-

siert werden. Dieser ist zwischen 18 und 30 Jahre alt und befindet sich noch in der Ausbildung an einer Hochschule. Da es sich meist um Studenten handelt, ist das Einkommen des Fernbusnutzers eher gering, sodass kein eigener PKW vorhanden ist, mit dem solche Strecken zurückgelegt werden könnten. Mit Hilfe der Conjoint-Analyse im Kapitel 4.3 können Merkmale aufgezeigt werden, die eine Entscheidung für die Lage der Fernbushaltestellen beeinflussen. Dabei zeigen sich für lange und kurze Strecken große Unterschiede. Bei langen Strecken, welche mehrere Städte anfahren, zeigt sich eine höhere Wichtigkeit für die Reisezeit. Die Kunden nehmen einen höheren Preis in Kauf, wenn sie dadurch eine große Zeitersparnis bekommen. So ist eine Haltestelle in Routennähe für lange Strecken eine sinnvolle Alternative, solange diese gut an das ÖPNV-Netz der jeweiligen Stadt angeschlossen ist. Dabei kann eventuell eine Kooperation zwischen Fernbusunternehmen und Verkehrsunternehmen bei Ticketpreisen denkbar sein. Bei kurzen Strecken, die keine zusätzlichen Städte anfahren, sind Haltestellen im Stadtzentrum besser geeignet. Hier kann durch eine veränderte Lage der Haltestelle meist keine Zeit eingespart werden. Bei der gemeinsamen Betrachtung zeigt sich wiederum, dass Haltestellen in Routennähe einen größeren Nutzen für Kunden aufweisen. Diese Information ist vor allem dadurch sehr wichtig, da die kurzen Strecken in fast allen Fällen Teilstrecken von langen Strecken sind. Es zeigt sich vor allem die Reisezeit als wichtiges Kriterium für die Fernbusreisenden. Es ist also vorstellbar, dass vor allem auf Strecken, die quer durch Deutschland verlaufen, durch Haltestellen in Routennähe mehr Kunden generiert werden können. Anschließend wurde der Einfluss der personenbezogenen Variablen auf die Wahlentscheidungen mittels logistischer Regressionsanalyse untersucht. Hierbei zeigt sich, dass diese Variablen auch einen Einfluss auf die Entscheidungen haben. Dieser ist jedoch eher gering, da Reisezeit und Reisekosten einen weitaus höheren Stellenwert haben. Dennoch können Geschlecht, Berufsstatus, Einkommen und PKW-Verfügbarkeit einen Einfluss darauf haben, für welche Variante sich die Person entscheidet.

5. Kritische Würdigung

Nach Abschluss der Analysen können die Ergebnisse kritisch betrachtet und hinterfragt werden. Zu Beginn können die „typischen“ Merkmale eines Fernbusnutzers mit einer Studie aus Kapitel 2 verglichen werden. Dabei zeigt sich eine gewisse Übereinstimmung der Resultate. In der Analyse in Kapitel 4 sind Fernbusnutzer zwischen 18 und 30 Jahre alt, gut gebildet und

besitzen meist keinen eigenen PKW. Auch in der Studie von Selzer, Kruse, Wilde und Lanzendorf (2016) sind viele Fernbusnutzer jünger als 35, besitzen Abitur oder einen Hochschulabschluss und haben keinen eigenen PKW. Die Nutzungshäufigkeit von Fernbussen ist ebenfalls in beiden Untersuchungen ähnlich. Allerdings kann in der Analyse in Kapitel 4 kein großer Unterschied zwischen Männern und Frauen festgestellt werden. Dies kann aber durchaus an der geringen Anzahl an Probanden liegen. Ein weiteres Merkmal, welches nicht untersucht worden ist, ist die Einstellung zum öffentlichen Personennahverkehr. Dies sollte in einer weiteren Arbeit erfolgen. Es muss allerdings auch betrachtet werden, dass sehr viel mehr Fernbusnutzer als Nicht-Nutzer vorhanden waren, sodass sich die Charakterisierung der Nicht-Nutzer eventuell mit einer größeren Anzahl an Probanden verändern könnte. In der darauffolgenden Conjoint-Analyse wurden die Auswahlentscheidungen bezüglich der Lage der Haltestellen untersucht. Hierbei waren vor allem Reisezeit und Reisekosten die unabhängigen Variablen. Allerdings ist für die Variante mit Haltestellen in Routennähe ein zusätzliches Ticket für den ÖPNV am Zielort in die Entscheidung eingegangen, welche nicht weiter in der Analyse betrachtet wurde, da dieses nur für Variante 2 zur Verfügung steht und somit keine Auswirkung auf die Ergebnisse gehabt hätte. Für eine weitere Arbeit sollten für beide Varianten verschiedene zusätzliche Tickets im ÖPNV betrachtet werden, um eine genauere Präferenz zu erkennen. Auch ist die mögliche Lage der Haltestelle nicht bekannt, sodass Zeiten zu und von Haltestellen geschätzt werden mussten. Es kann also vorkommen, dass es Kunden gibt, die sehr nah an neuen Haltestellen in Routennähe wohnen, sodass auch für kurze Strecken Zeit eingespart werden könnte. Eventuell könnten in nachfolgenden Arbeiten verschiedene Wohnorte in Städten mit kurzer, mittlerer und langer Reisezeit zu und von den Haltestellen betrachtet werden.

6. Zusammenfassung und Ausblick

Durch die angewendeten Analysemethoden können verschiedene Ergebnisse zu den Merkmalen von Fernbusnutzern, zur Lage der Haltestelle und zu Einflüssen auf die Auswahlentscheidungen von Kunden von Fernbusstrecken gezogen werden. Ein „typischer“ Fernbusnutzer ist zwischen 18 und 30 Jahre alt und gehört somit zu der jüngeren Generation. Der Großteil besitzt noch keine abgeschlossene Berufsausbildung und befindet sich noch im Studium. Dadurch ist das Einkommen der Kunden eher gering, welches für die Nutzung eines Fernbusses

spricht, da dieser eine günstige Reisealternative darstellt. Ein wichtiger Punkt ist, dass Fernbusnutzer nur in wenigen Fällen einen eigenen PKW besitzen. Dies spricht ebenfalls für eine Nutzung des Fernbusses, welcher von den meisten Kunden mehrmals im Jahr genutzt wird. In der anschließenden Conjoint-Analyse konnten Erkenntnisse über die optimale Lage der Haltestellen aus Kundensicht gewonnen werden. Dabei stellt sich heraus, dass für lange Strecken mit mehreren Zwischenhalten, die Haltestellen des Fernbusses eher in Routennähe liegen sollten, da somit sehr viel Zeit eingespart wird. Wichtig ist dabei eine gute Anbindung zum ÖPNV. Bei kurzen Strecken sind Haltestellen in der Stadt eine bessere Alternative. Da aber viele kurze Strecken ein Teil von langen Strecken sind, wurden ebenfalls alle Strecken zusammen betrachtet. Auch hier zeigt sich, dass Haltestellen in Routennähe einen höheren Nutzen für die Kunden darstellen. Der ausschlaggebende Einfluss auf die Entscheidung stellt dabei die Reisezeit dar. Durch die logistische Regression konnte zusätzlich festgestellt werden, dass die personenspezifischen Merkmale keinen großen Einfluss auf die Auswahlentscheidungen besitzen. Zusätzliche Untersuchungen könnten sich mit Kosten-Nutzen-Analysen aus Unternehmenssicht beschäftigen. Des Weiteren können eventuell noch umfangreichere Umfragen zur Lage von Haltestellen durchgeführt werden. Dabei sollten die Fahrzeiten im ÖPNV in den verschiedenen Städten variiert werden, um unterschiedliche Wohnorte zu simulieren.

Quellenverzeichnis

- Akaike, H. (1973), *Information Theory and an Extension of the Maximum Likelihood Principle*, Biometrika Vol. 60, No. 2 (Aug., 1973), S. 255-265
- Augustin, K., Gerike, R., Sanchez, M.J.M, Ayala,C. (2014), *Analysis of intercity bus markets on long distance in an establishes and a young market: The example of the U.S. and Germany*, Research in Transportation Economics 48, S. 245-254
- Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2011), *Multivariate Analysemethoden- Eine anwendungsorientierte Einführung*, 13. Auflage, Springer, Berlin, Heidelberg
- Backhaus, K., Erichson, B., Plinke, W., Weiber, R. (2016), *Multivariate Analysemethoden- Eine anwendungsorientierte Einführung*, 14. Auflage, Springer, Berlin, Heidelberg
- Backhaus, K., Erichson, B., Weiber, R. (2015), *Fortgeschrittene Multivariate Analysemethoden – Eine anwendungsorientierte Einführung*, 3. Auflage, Springer, Berlin, Heidelberg
- Bange, C. (2013), „Planung, Finanzierung und Betrieb von Fernbusterminals in Deutschland“, Karlsruher Institut für Technologie, Vortrag, Konferenz „Kommunales Infrastruktur-Management“
- Böckers, V., Haucamp, J., Heimeshoff, U., Thorwarth, S. (2015), *Auswirkungen der Fernbusliberalisierung auf den Schienenpersonenverkehr*, List Forum 41, S. 75-90
- Fabian, S. (2005), *Wettbewerbsforschung und Conjoint-Analyse*, Wiesbaden
- Fitzner, J., Schmidt, H.J. (2016), *Fernbuslinien im Fokus: Eine Conjoint-Analyse zur Bestimmung der kaufrelevanten Produkteigenschaften*, Wissenschaftliche Schriften des Fachbereichs Wirtschaftswissenschaften, Hochschule Koblenz – University of Applied Sciences, No.22
- FlixBus (2014), *Studie zur Kundenfreundlichkeit der deutschen Fernbusbahnhöfe*, in Zusammenarbeit mit der deutschen Hochschule Heilbronn, Präsentation
- Knorr, A., Lueg-Arndt, A. (2016), *Intercity bus deregulation in Germany – Intramodal and intermodal effect after two years*, Research in Transportation Economics 59, S. 323-329

Mäertens, S. (2012), *Buslinienverkehr in Deutschland – effiziente Ausgestaltung einer Liberalisierung*, Wirtschaftsdienst, Analysen und Berichte

MeinFernbus, Flixbus, Streckennetz und Fahrpreise, <https://meinfernbus.de/> (Stand 24.03.2017)

Moonroc Institute of Economic Research (2013), *Der Fernbus – Mobilität mit Zukunft oder nur rein überschätzter Hype?*, Artikelnummer 2013007

Mosler, K., Schmid, F. (2009), *Beschreibende und Wirtschaftsstatistik*, 4.Auflage, Springer, Heidelberg

Selzer, S., Kruse, K., Wilde, M., Lanzendorf, M. (2016), *Integration von Fernbuslinienangeboten. Anforderungen an und Handlungsoptionen für die städtebauliche und verkehrliche Integration der Fernbusse in lokale Verkehrssysteme. Ergebnisse einer Fahrgastbefragung in Frankfurt am Main*, Arbeitspapiere zur Mobilitätsforschung Nr. 13



Erklärung zur Urheberschaft

Ehrenwörtliche Erklärung

Hiermit erkläre ich ehrenwörtlich, dass ich die vorliegende Arbeit selbst angefertigt habe. Die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher noch keiner anderen Prüfbehörde vorgelegt und noch nicht veröffentlicht. Ich bin mir bewusst, dass eine unwahre Erklärung rechtliche Folgen haben wird.

Dresden, den 08.08.2017

Jana-Christina Hambrock

Anhang

A.1 Fragebogen

Angaben zur Person

- **Geschlecht** ☐ männlich ☐ weiblich

- **Alter**
☐ 18 – 25 ☐ 26 – 30 ☐ 31 – 35 ☐ 36 – 45 ☐ 46 – 55 ☐ 56 – 64 ☐ 65+

- **Beruf**
☐ Schüler ☐ Student ☐ Azubi ☐ Berufstätig ☐ Rentner ☐ Arbeitssuchend

- **Monatliches Einkommen**
☐ < 1000€ ☐ >1000€ – 2000€ ☐ >2000€ - 3000€ ☐ >3000€ - 4500€

☐ >4500€

- **Höchster Schulabschluss**
☐ Hauptschule ☐ Realschule ☐ Abitur ☐ ohne Schulabschluss

- **Höchste Berufsausbildung**
☐ Lehre / Berufsfachschule ☐ Meister-/Fachschule
☐ Hoch-/Fachhochschulabschluss ☐ (noch) ohne Berufsausbildung

- **Eigener PKW vorhanden?** ☐ Ja ☐ Nein

- **Sind Sie schon einmal mit einem Fernbus gefahren?** ☐ Ja ☐ Nein

- **Wenn ja, wie oft nutzen Sie einen Fernbus?**
☐ 1-2x im Jahr ☐ 1-2x in 6 Monaten ☐ 1-2x in 3 Monaten ☐ 1-2x im Monat

☐ 1-2x in der Woche



Conjoint-Fragen

Im folgenden Teil des Fragebogens werden für verschiedene Streckenverbindungen mit dem Fernbus zwei Wahlmöglichkeiten vorgestellt.

Bitte stellen Sie sich vor, Sie würden in einer Stadt wohnen und den Fernbus für die jeweilige Strecke nutzen. Bitte entscheiden Sie sich für die Variante, die Sie bevorzugen würden, unabhängig davon, ob Sie den Fernbus in Wirklichkeit für diese Strecke wählen würden oder nicht.

Es stehen Ihnen keine zusätzlichen ÖV-Tickets (Job-Ticket, Semesterticket,...) zur Verfügung. Sie müssen sich also für die Strecken vor und nach dem Fernbus ein extra Ticket kaufen. Ausnahme ist nur das in manchen Varianten angegebene ÖV-Ticket. Dieses gilt nur für Strecke von Haltestelle bis Zielpunkt nach der Fahrt im Fernbus.

1. Fall: Strecke Dresden - Hannover

1) Reisezeit im Bus: 6:20h

Kosten: 16€

Reisezeit Startpunkt bis Haltestelle: 20min

☐

Reisezeit Haltestelle bis Zielpunkt: 25min

Komplexe Reisezeit: 7:05h

Kein ÖV-Ticket

2) Reisezeit im Bus: 4:50h

Kosten: 19€

Reisezeit Startpunkt bis Haltestelle: 35min

☐

Reisezeit Haltestelle bis Zielpunkt: 40min

Komplexe Reisezeit: 6:05h

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

2. Fall: Strecke Hamburg – München

1) Reisezeit im Bus: 12:45h

Kosten: 30€



Reisezeit Startpunkt bis Haltestelle: 15min

Reisezeit Haltestelle bis Zielpunkt: 25min

Komplexe Reisezeit: 13:25h

Kein ÖV-Ticket

2) Reisezeit im Bus: 9:45h

Kosten: 33€



Reisezeit Startpunkt bis Haltestelle: 30min

Reisezeit Haltestelle bis Zielpunkt: 40min

Komplexe Reisezeit: 10:55h

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

3. Fall: Dresden – München

1) Reisezeit im Bus: 8:25h

Kosten: 21€



Reisezeit Startpunkt bis Haltestelle: 20min

Reisezeit Haltestelle bis Zielpunkt: 25min

Komplexe Reisezeit: 9:10h

Kein ÖV-Ticket

2) Reisezeit im Bus: 6:55h

Kosten: 24 €

Reisezeit Startpunkt bis Haltestelle: 35min



Reisezeit Haltestelle bis Zielpunkt: 40min

Komplexe Reisezeit: 8:10h

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

4. Fall: Berlin – Frankfurt

- 1) Reisezeit im Bus: 11:40

Kosten: 25€

Reisezeit Startpunkt bis Haltestelle: 30min



Reisezeit Haltestelle bis Zielpunkt: 20min

Komplexe Reisezeit: 12:30h

Kein ÖV-Ticket

- 2) Reisezeit im Bus: 8:10

Kosten: 28€

Reisezeit Startpunkt bis Haltestelle: 40min



Reisezeit Haltestelle bis Zielpunkt: 35min

Komplexe Reisezeit: 9:25h

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

5. Fall: Dresden – Leipzig

- 1) Reisezeit im Bus: 1:50h

Kosten: 5,50€

Reisezeit Startpunkt bis Haltestelle: 20min



Reisezeit Haltestelle bis Zielpunkt: 15min

Komplexe Reisezeit: 2:25h

Kein ÖV-Ticket

- 2) Reisezeit im Bus: 1:50h

Kosten: 8,50€

Reisezeit Startpunkt bis Haltestelle: 35min



Reisezeit Haltestelle bis Zielpunkt: 20min

Komplexe Reisezeit: 2:45min

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

6. Fall: Hamburg – Hannover

- 1) Reisezeit im Bus: 2:10h
Kosten: 10€
Reisezeit Startpunkt bis Haltestelle: 15min
Reisezeit Haltestelle bis Zielpunkt: 25min
Komplexe Reisezeit: 2:50min
Kein ÖV-Ticket ☐
- 2) Reisezeit im Bus: 2:10h
Kosten: 13€
Reisezeit Startpunkt bis Haltestelle: 30min
Reisezeit Haltestelle bis Zielpunkt: 40min
Komplexe Reisezeit: 3:20h
ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig) ☐

7. Fall: Berlin – Leipzig

- 1) Reisezeit im Bus: 2:10h
Kosten: 8€
Reisezeit Startpunkt bis Haltestelle: 30min
Reisezeit Haltestelle bis Zielpunkt: 15min
Komplexe Reisezeit: 2:55h
Kein ÖV-Ticket ☐
- 2) Reisezeit im Bus: 2:10h
Kosten: 11€
Reisezeit Startpunkt bis Haltestelle: 40min
Reisezeit Haltestelle bis Zielpunkt: 20min
Komplexe Reisezeit: 3:10min
ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig) ☐



8. Fall: München – Nürnberg

1) Reisezeit im Bus: 2:35min

Kosten: 6€

Reisezeit Startpunkt bis Haltestelle: 25min



Reisezeit Haltestelle bis Zielpunkt: 20min

Komplexe Reisezeit: 3:20h

Kein ÖV-Ticket

2) Reisezeit im Bus: 2:35min

Kosten: 9€

Reisezeit Startpunkt bis Haltestelle: 40min



Reisezeit Haltestelle bis Zielpunkt: 30min

Komplexe Reisezeit: 3:45min

ÖV-Ticket im Preis enthalten (In Zielstadt eine Einzelfahrt, zonenunabhängig)

A.2 Zwischenstopps pro Strecke

Strecke	Zwischenstopps
Dresden – Hannover	Braunschweig, Magdeburg, Leipzig
Hamburg – München	Hannover, Göttingen, Kassel, Fulda, Schweinfurt, Nürnberg
Dresden – München	Prag, Pilsen, Regensburg
Berlin – Frankfurt am Main	Magdeburg, Braunschweig, Göttingen, Kassel, Bad Zwesten, Marburg, Giessen

Quelle: MeinFernbus Flixbus www.meinfernbus.de Stand: 24.03.2017



A.3 Korrelationsmatrix

	Geschlecht	Alter	Beruf	Einkommen	Schulabschluss	Berufsausbildung	PKW vorhanden	Fernbus gefahren	Wie oft?
Geschlecht	1								
Alter	0,05667029	1							
Beruf	-0,04547289	0,63826194	1						
Einkommen	0,17765619	0,50043818	0,56433449	1					
Schulabschluss	-0,07865665	-0,52817501	-0,37716471	-0,26148383	1				
Berufsausbildung	-0,01854397	-0,40186126	-0,53988113	-0,40031248	0,424412918	1			
PKW vorhanden	-0,00828941	0,42616938	0,46355006	0,45083196	-0,387298335	-0,360470211	1		
Fernbus gefahren	-0,09107291	-0,46156786	-0,40834565	-0,30860583	0,310238357	0,299371519	-0,460593396	1	
Wie oft?	0,17690156	0,41627565	0,3921354	0,33599498	-0,279040169	-0,274931016	0,421393178	-0,829976938	1



Digitale Fassung der Masterarbeit auf CD-ROM

Enthält folgende Dateien:

- Text der Masterarbeit (.pdf)
- Text der Masterarbeit (.doc)
- Tabellen (.xls)
- Zitierte Studien (.pdf)
- R-Code
- .csv-Dateien für den R-Code
- Erstellte Grafiken (.png)
- Fragebogen (.doc)